

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2002, 10:22:17 ; Search time 67 seconds

(without alignments)
920.821 Million cell updates/sec

Title: US-09-816-467-2

Perfect score: 2456

Sequence: 1 MFSPFIPPSYSKNLDCWVD.....DKILGCDWYFPTDEGTND 463

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2456	100.0	463	20	AAV00921
2	2451	99.8	573	8	AAV07045
3	2444	99.5	1315	22	AAV07169
4	2440	99.3	618	19	AAV08909
5	2397	97.6	665	22	AAE07895
6	2397	97.6	882	22	AAE07889
7	2397	97.6	907	22	AAE07891
8	2397	97.6	1052	22	AAE07903
9	2397	97.6	1112	22	AAE07902
10	2393	97.4	452	12	AAV12471

11	2387.5	97.2	661	22	AAV01429
12	2385	97.1	453	22	AAV01427
13	2073	84.4	605	22	AAE07897
14	1578.5	64.3	660	22	AAE07898
15	1261	51.3	448	22	AAE07896
16	653.5	26.6	1067	21	AAV93307
17	653.5	26.6	1092	21	AAV93310
18	653.5	26.6	1296	17	AAV95010
19	651.5	26.5	1295	23	AAV95010
20	642.5	26.2	449	21	AAV77139
21	642.5	26.2	449	22	AAV04167
22	642.5	26.2	473	21	AAV68400
23	642.5	26.2	837	21	AAV77140
24	630.5	25.7	847	22	AAV04081
25	627.5	25.5	434	22	AAV04089
26	627.5	25.5	435	22	AAV04090
27	627.5	25.5	437	22	AAV04088
28	627.5	25.5	438	17	AAV95008
29	627.5	25.5	438	19	AAV68389
30	627.5	25.5	438	21	AAV77134
31	627.5	25.5	445	19	AAV68391
32	627.5	25.5	462	17	AAV95009
33	627.5	25.5	462	19	AAV68390
34	616.5	25.1	432	21	AAV77142
35	611	24.9	415	22	AAV04083
36	607	24.7	472	19	AAV68393
37	603	24.6	382	21	AAV68393
38	596	24.3	439	22	AAV04085
39	594	24.2	440	22	AAV77135
40	594	24.2	440	22	AAV04091
41	592	24.1	1291	19	AAV68392
42	590	24.0	472	19	AAV68394
43	590	24.0	1070	21	AAV93308
44	590	24.0	1095	21	AAV93311
45	573.5	23.4	452	19	AAV68396

ALIGNMENTS

RESULT 1
ID: AAV00921 standard: Protein: 463 AA.
AC: AAV00921;
XX: 28-MAY-1999 (first entry)
XX: Tetanus toxin fragment C protein sequence.
XX: Tetanus toxin fragment C; TTC; central nervous system; CNS; spinal cord; proteolytic fragment; retrograde axonal transport; spinal cord disease; transsynaptic transport; neurodegenerative disease; motoneuron disease; amyotrophic lateral sclerosis; spinal muscular atrophy; therapy; ALS; SMA; neurodegenerative lysosomal storage disease; neuronal mapping.
XX: Clostridium tetani.
XX: MO9909057-A2.
XX: 25-FEB-1999.
XX: 12-AUG-1998; 98MO-EP05113.
XX: 13-NOV-1997; 97US-0065236.
XX: 14-AUG-1997; 97US-0055615.
XX: (INSP) INST PASTEUR.
XX: Brulet P, Coen L, Osta Pinzolas R;
XX: WPI: 1999-180971/15.
XX: N-PSDB; AAX27234.

XX Delivery of a composition to the central nervous system or spinal
 PT cord - comprises administration of a non-toxic, proteolytic
 PT fragment of tetanus toxin in association with a molecule having
 PT biological function

XX Example 1; Fig 1; 53pp: English.

XX This sequence represents the tetanus toxin fragment C (TTC).
 CC The invention relates to a method for in vivo delivery of a desired
 CC composition into a human or animal central nervous system (CNS) or spinal
 CC cord comprising administering a non-toxic, proteolytic fragment of tetanus
 CC toxin (TT) in association with at least a molecule having a biological
 CC function and where the composition is capable of in vivo retrograde
 CC axonal transport and transsynaptic transport into the CNS or the spinal
 CC cord of the human or animal and of being delivered to different areas of
 CC the CNS or the spinal cord. The method can be used for the treatment of
 CC humans or animals with CNS or spinal cord disease, e.g. neurodegenerative
 CC and motoneuron diseases such as amyotrophic lateral sclerosis (ALS),
 CC spinal muscular atrophies (SMA) or neurodegenerative lysosomal storage
 CC diseases. Compositions comprising hybrid fragments of TT comprising
 CC fragments C and B can also be used for neuronal mapping and
 CC immunisations. Use of TT comprising fragments A, B and C results in
 CC better transport of the fragment inside the organism compared with
 CC fragment C.

XX Sequence 463 AA;

Query Match Best Local Similarity 100.0%; Score 2456; DB 20; Length 463;

Matches 463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSPTPIPSYSKNDKCDWNEEDIDVILKSTIINDNDIISDGFNSVITTPDA 60
 Db 1 MFSPTPIPSYSKNDKCDWNEEDIDVILKSTIINDNDIISDGFNSVITTPDA 60
 QY 61 QLVPTINKAHLVNNESSEVIVHKAMDEYNDMFNFVSWLRVPKVSASHLEQYGTN 120
 Db 61 QLVPTINKAHLVNNESSEVIVHKAMDEYNDMFNFVSWLRVPKVSASHLEQYGTN 120
 QY 121 EYSIISSMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKFNAYLANKWV 180
 Db 121 EYSIISSMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKFNAYLANKWV 180
 QY 181 FTTTNDRLSSANLYINGVLMGSAEITGGAIREDDNNTTLKLDRCNNNNQYVSTDKFRIFC 240
 Db 181 FTTTNDRLSSANLYINGVLMGSAEITGGAIREDDNNTTLKLDRCNNNNQYVSTDKFRIFC 240
 QY 241 CKALNPKEIEKLYTSYSTIFLDFMGNPLRYDTEYVLIIPVASSSKDVOLKNTIDYMYLT 300
 Db 241 CKALNPKEIEKLYTSYSTIFLDFMGNPLRYDTEYVLIIPVASSSKDVOLKNTIDYMYLT 300
 QY 301 NAPSITNGKLNITYRRLYNGLKFIIKRYTPNNEIDSFVKSQGFITKLYSYNNNEHIVGYR 360
 Db 301 NAPSITNGKLNITYRRLYNGLKFIIKRYTPNNEIDSFVKSQGFITKLYSYNNNEHIVGYR 360
 QY 361 KGNNAFNNNDRLIRVGYNAAGIPLKYMKAIVKLDLKYTSVQKLKYDDKNAASGLVGTGTHG 420
 Db 361 KGNNAFNNNDRLIRVGYNAAGIPLKYMKAIVKLDLKYTSVQKLKYDDKNAASGLVGTGTHG 420
 QY 421 GOIGNDPNRDIILASNMVFNHLKDKIIGCDWYFVPTDEGWTND 463
 Db 421 GOIGNDPNRDIILASNMVFNHLKDKIIGCDWYFVPTDEGWTND 463

RESULT 2
 ID AAF70345 standard; Protein; 573 AA.

XX AAF70345;
 AC
 XX
 DT 22-APR-1991 (first entry)
 XX

DE Portion of B fragment and all of the C fragment of tetanus toxin.

XX TT; vaccine.

XX Clostridium tetani.

XX EP209281-A.

XX 21-JAN-1987.

XX 27-JUN-1986; 86EP-0305029.

XX 28-JUN-1985; 85GB-0016442.

XX (WELL) WELLCOME FOUNDATION LTD.

XX WPI: 1987-015999/03.

XX N-PSDB; AAN70545.

XX Claim 4; Fig 1; 36pp: English.

XX Gene product comprises a tetanus toxin fragment, which may be

XX expressed in a transformed host, and used as an antigen in vaccine

XX production, against the disease.

XX Sequence 573 AA;

Query Match Best Local Similarity 99.8%; Score 2451; DB 8; Length 573;

Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VFSPTPIPSYSKNDKCDWNEEDIDVILKSTIINDNDIISDGFNSVITTPDAQ 61
 Db 112 VFSPTPIPSYSKNDKCDWNEEDIDVILKSTIINDNDIISDGFNSVITTPDAQ 171
 QY 62 LVPINGKAIHLVNNESSEVIVHKAMDEYNDMFNFVSWLRVPKVSASHLEQYGTNE 121
 Db 172 LVPINGKAIHLVNNESSEVIVHKAMDEYNDMFNFVSWLRVPKVSASHLEQYGTNE 231
 QY 122 YSIISMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKFNAYLANKWV 181
 Db 223 YSIISMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKFNAYLANKWV 291
 QY 182 ITTNDRLSSANLYINGVLMGSAEITGGAIREDDNNTTLKLDRCNNNNQYVSTDKFRIFC 241
 Db 292 ITTNDRLSSANLYINGVLMGSAEITGGAIREDDNNTTLKLDRCNNNNQYVSTDKFRIFC 351
 QY 242 KALNPKEIEKLYTSYSTIFLDFMGNPLRYDTEYVLIIPVASSSKDVOLKNTIDYMYLTN 301
 Db 352 KALNPKEIEKLYTSYSTIFLDFMGNPLRYDTEYVLIIPVASSSKDVOLKNTIDYMYLTN 411
 QY 302 APSYTGKLNITYRRLYNGLKFIIKRYTPNNEIDSFVKSQGFITKLYSYNNNEHIVGYR 361
 Db 412 APSYTGKLNITYRRLYNGLKFIIKRYTPNNEIDSFVKSQGFITKLYSYNNNEHIVGYR 471
 QY 362 DGNNAFNNNDRLIRVGYNAAGIPLKYMKAIVKLDLKYTSVQKLKYDDKNAASGLVGTGTHG 421
 Db 472 DGNNAFNNNDRLIRVGYNAAGIPLKYMKAIVKLDLKYTSVQKLKYDDKNAASGLVGTGTHG 531
 QY 422 OTGNDPNRDIILASNMVFNHLKDKIIGCDWYFVPTDEGWTND 463
 Db 532 OTGNDPNRDIILASNMVFNHLKDKIIGCDWYFVPTDEGWTND 573

RESULT 3
 ID AAB61169 standard; Protein; 1315 AA.

XX

AC AAB61169;
 XX
 DF 02-APR-2001 (first entry)
 XX
 DE Clostridium tetani TeNT.
 XX
 KM Clostridium tetani: TeNT; tetanus toxin; antibacterial; vaccine;
 KM TeNT fragment C; infection.
 XX
 OS Clostridium tetani.
 XX
 PN W0200100839-A1.
 PD 04-JAN-2001.
 XX
 PF 23-JUN-2000; 2000WO-GH02428.
 XX
 PR 25-JUN-1999; 99GB-0014861.
 XX
 PA (UNLO) IMPERIAL COLLEGE SCI TECHNOLOGY & MED.
 XX
 PI Fairweather NF, Sinha K;
 XX
 DR WPI; 2001-123014/13.
 XX
 PT New polypeptide, useful for treating infections of Clostridium tetani,
 XX comprises tetanus toxin fragment with a mutation in a loop region,
 PS
 XX
 XX Disclosure: Page 39; 43pp; English.
 CC The present sequence is given in a specification relating to a novel
 CC polypeptide comprising tetanus toxin (TeNT) fragment C or its immunogenic
 CC fragment, containing a mutation in a loop region. The mutation results in
 CC a reduction in the binding of TeNT fragment G or its immunogenic fragment
 CC to gangliosides and primary motoneurons, and/or a reduction in the
 CC ability of TeNT fragment C or its immunogenic fragment to undergo
 CC retrograde transport. The polypeptide is useful for treating, preventing
 CC and reducing the susceptibility to Clostridium tetani infection in a
 CC human or animal, and also for producing antibodies which recognise groups
 CC within TeNT polypeptides. Antibody produced against the polypeptide is
 CC also useful for treating Clostridium tetani infection.
 XX
 SQ Sequence 1315 AA;
 Query Match 99.5%; Score 2444; DB 22; Length 1315;
 Best Local Similarity 99.8%; Pred. No. 4.2e-189;
 Matches 461; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 VESTPIPSYSKNIIDCWVNEEDIDVILKSTIILNDIINDIISDGSFNSVITYPDAQ 61
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 854 VESTPIPSYSKNIIDCWVNEEDIDVILKSTIILNDIINDIISDGSFNSVITYPDAQ 913
 QY 62 LVPGINGKAHLVNNESSEVIVHKAMDIENYDMFNFTVFWLVPKVSASHLOXYGNE 121
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 914 LVPGINGKAHLVNNESSEVIVHKAMDIENYDMFNFTVFWLVPKVSASHLOXYGNE 973
 QY 122 YSIISMKKHSLSIGSGWSVSLKGNLNIWTLKDSAGEVROITFRDLPDKENAYLANKWVF 181
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 974 YSIISMKKHSLSIGSGWSVSLKGNLNIWTLKDSAGEVROITFRDLPDKENAYLANKWVF 1033
 QY 182 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNTTLKDRNNNNNOYVSDIKFRIFC 241
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1034 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNTTLKDRNNNNNOYVSDIKFRIFC 1093
 QY 242 KALNPKRIEKLITYSITFLRPFWGNPLRYDEYLLIPVASSSKDQVLKNTIDYMLTN 301
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1094 KALNPKRIEKLITYSITFLRPFWGNPLRYDEYLLIPVASSSKDQVLKNTIDYMLTN 1153
 QY 302 APSYTGKILNIYYRRLYNGIKFKITIRYTPNNEIDSFVKSQDFIKLYVSYNNNHVIGYPR 361
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1154 APSYTGKILNIYYRRLYNGIKFKITIRYTPNNEIDSFVKSQDFIKLYVSYNNNHVIGYPR 1213
 QY 362 DGNAFNNLDRILRVGYNABGIPLYKKMEAVKLRDLKTSYVOLKLYDDKNASLGLVGTGNG 421

DB 1214 DGNAFNNLDRILRVGYNABGIPLYKKMEAVKLRDLKTSYVOLKLYDDKNASLGLVGTGNG 1273
 QY 422 QIGNDPNRDIILASNMWYNHKLKDKITLGCDFWFPVPTDEGTND 463
 DB 1274 QIGNDPNRDIILASNMWYNHKLKDKITLGCDFWFPVPTDEGTND 1315
 RESULT 4
 AAW48909
 ID AAW48909 standard; Protein: 618 AA.
 XX
 AC AAW48909;
 XX
 DT 23-SEP-1998 (first entry)
 XX
 DE SOD-1/TTC hybrid protein.
 XX
 KM Chimeric: copper-zinc superoxide dismutase; SOD-1; TTC; SOD:Te1451;
 KM tetanus toxin fragment C; tetanus holotoxin; nerve cell; stroke;
 KM neurological disorder; oxidative stress; brain hypoxia-reperfusion;
 XX epilepsy; Parkinson's disease; Huntington's disease.
 OS Chimeric - Homo sapiens.
 XX
 OS Chimeric - Clostridium tetani.
 XX
 FH Key Location/Qualifiers
 FT Region 1..163
 FT Region /note="SOD-1"
 FT Region 168..618
 FT Region /note="TTC moiety"
 XX
 XX US5780024-A.
 XX
 XX 14-JUL-1998.
 XX
 XX 21-JUN-1996; 96US-0668381.
 XX
 XX 23-JUN-1995; 95US-0000473.
 XX
 XX 21-JUN-1996; 96US-0668381.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Brown RH, Fishman PS, Francis JW, Hosler BA;
 DR WPI; 1998-412999/35.
 DR N-PSDB; AAV32580.
 XX
 PT New hybrid protein of superoxide dismutase and tetanus toxin
 PT fragment C - having increased uptake by neurons and retention of
 PT enzymatic activity in these cells, for treating neurological
 PT diseases associated with oxidative stress
 XX
 PS Claim 7; Columns 23-26; 23pp; English.
 XX
 CC The present sequence represents an enzymatically active human
 CC copper-zinc superoxide dismutase (SOD-1) fused at its carboxyl
 CC terminus with the tetanus toxin fragment C (TTC) moiety. The TTC
 CC moiety constitutes amino acid residues 865-1315 of the tetanus
 CC holotoxin. The hybrid protein, referred as SOD:Te1451, is claimed
 CC to have the following properties: (a) it exhibits Cu/Zn SOD enzymatic
 CC activity; (b) the TTC moiety selectively binds to nerve cells and
 CC allows uptake of the hybrid protein into these cells; and (c) it
 CC retains substantial SOD enzymatic activity following cellular uptake.
 CC SOD:Te1451 is claimed to be useful for treating neurological disorders
 CC associated with oxidative stress, e.g. stroke, brain hypoxia-reperfusion,
 CC epilepsy, Parkinson's and Huntington's diseases.
 XX
 SO Sequence 618 AA;
 Query Match 99.3%; Score 2440; DB 19; Length 618;
 Best Local Similarity 99.6%; Pred. No. 3e-189;

CC The invention relates to a non toxic polypeptide, for delivery of a
CC therapeutic agent to a neuronal cell, which comprises a binding domain
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC as Hc) that binds to the neuronal cell and a translocation domain (amino
CC terminal half of HC, designated as HN), that translocates the therapeutic
CC agent into the neuronal cell, where the translocation domain is not a HN
CC domain of a clostridial neurotoxin and is not a fragment or derivative of
CC a HN domain of a clostridial toxin. Polypeptides of the invention are
CC useful for the treatment of a disease state associated with neuronal
CC cells. The polypeptide constructs are useful for delivering therapeutic
CC substances to neuronal cells. They are useful to treat disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is modified clostridial heavy chain fragment. This sequence is
CC constructed by fusing the binding domain of tetanus neurotoxin (TcNT)
CC with translocation domain of diphtheria neurotoxin.

SD Sequence 665 AA;

QY Query Match 97.6%; Score 2397; DB 22; Length 665;
Best Local Similarity 100.0%; Pred. No. 1e-185;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 12 SKNLDQWVNEEDIVYLKKSTLLNDINDNDITSDISGENSSVIYYPDAOLVPGINGKAI 71
|||||
Db 212 SKNLDQWVNEEDIVYLKKSTLLNDINDNDITSDISGENSSVIYYPDAOLVPGINGKAI 271

QY 72 HLVNNSSESEVIYHAKMDIEYNDFNNFVYSFYLRYVKVSAASHLEOYGTNETSIISMKKH 131
|||||
Db 272 HLVNNSSESEVIYHAKMDIEYNDFNNFVYSFYLRYVKVSAASHLEOYGTNETSIISMKKH 331

QY 132 SLSTSGWWSYLLKGNLLIWTLLKDSAGEVRIRITPRDLDPDKENAVLANKKWFEITITNDRLLS 191
|||||
Db 332 SLSTSGWWSYLLKGNLLIWTLLKDSAGEVRIRITPRDLDPDKENAVLANKKWFEITITNDRLLS 391

Query Match	Best Local Similarity	Matches	Conservative	Score	DB	Length
12	97.6%	100.0%	0	2397	DB 22	665
212	100.0%	0	0	1e-185		
72	100.0%	0	0			
272	100.0%	0	0			
132	100.0%	0	0			
332	100.0%	0	0			
192	100.0%	0	0			
392	100.0%	0	0			
252	100.0%	0	0			
452	100.0%	0	0			
312	100.0%	0	0			
512	100.0%	0	0			
372	100.0%	0	0			
572	100.0%	0	0			
432	100.0%	0	0			
632	100.0%	0	0			
AAE07889	100.0%	0	0			
AAE07889	100.0%	0	0			
01-NOV-2001	100.0%	0	0			
Modified clostridial heavy chain-superoxide dismutase conjugate #1.						
Neuronal cell; binding domain; translocation domain; stroke; epilepsy						
tumour; infection; neurodegenerative disease; gene therapy; chimeric;						
superoxide dismutase; SOD; diphtheria neurotoxin; tetanus neurotoxin;						
TENT.						
Chimeric - Bacillus stearothermophilus.						
Chimeric - Corynebacterium diphtheriae.						

Query Match 97.6%; Score 2397; DB 22; Length 907;
 Best Local Similarity 100.0%; Pred. No. 1.6e-185;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 SKNLDGWNEEDIDVILKSTILNDINNDIISDISGNSSVITTPDAQLVPGINGKAI 71
 DB 454 SKNLDGWNEEDIDVILKSTILNDINNDIISDISGNSSVITTPDAQLVPGINGKAI 513
 QY 72 HLNNSSSEVIYKRAMDIEYNDMFNFTVFWLRVPKVSASHLEOYGTNEYSIISSMKH 131
 DB 514 HLNNSSSEVIYKRAMDIEYNDMFNFTVFWLRVPKVSASHLEOYGTNEYSIISSMKH 573
 QY 132 SLISGGSWSLKGNNLIWTLKDSAGEVROITPDLDPKFNAYLANKWVFITITNDRLSS 191
 DB 574 SLISGGSWSLKGNNLIWTLKDSAGEVROITPDLDPKFNAYLANKWVFITITNDRLSS 633
 QY 192 ANLYINGVLMSAEITGLGAIREDNNTTLKLDRCNNNNQYVSIDKFRIEFCALNPKIEK 251
 DB 634 ANLYINGVLMSAEITGLGAIREDNNTTLKLDRCNNNNQYVSIDKFRIEFCALNPKIEK 693
 QY 252 LKYSYSLITPLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNTTMYLTNAPSTYNGKLN 311
 DB 694 LKYSYSLITPLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNTTMYLTNAPSTYNGKLN 753
 QY 312 IYRRLYNGLKFTIKRTPPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPKGNAFNNIDR 371
 DB 754 IYRRLYNGLKFTIKRTPPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPKGNAFNNIDR 813
 QY 372 ILRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNGQIGNDPNRI 431
 DB 814 ILRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNGQIGNDPNRI 873
 QY 432 LIASNMYFNHLKDKILGCDWYFVPTDEGWTND 463
 DB 874 LIASNMYFNHLKDKILGCDWYFVPTDEGWTND 905

RESULT 8
 AAE07903
 ID AAE07903 standard; Protein; 1052 AA.
 AC AAE07903;
 DT 01-NOV-2001 (first entry)
 XX
 C. botulinum C2 translocation domain with TeNT binding domain #2.
 DE
 XX Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
 KW tumour; infection; neurodegenerative disease; gene therapy;
 KM botulinum neurotoxin; tetanus neurotoxin; TeNT.
 XX
 OS Clostridium botulinum.
 OS Clostridium tetani.
 XX
 PN MO200158936-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 04-DEC-2000; 2000MO-GB04644.
 XX
 PR 02-DEC-1999; 99GB-0028530.
 PR 07-APR-2000; 2000GB-0008658.
 XX
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
 XX
 PI Shone CC, Sutton JM, Silman N;
 XX
 DR WPT; 2001-514643/56.
 XX
 XX New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -
 XX

PS Example 2; Page 50; 50bp; English.
 XX The invention relates to a non toxic polypeptide, for delivery of a
 CC therapeutic agent to a neuronal cell, which comprises a binding domain
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
 CC as HC) that binds to the neuronal cell and a translocation domain (amino
 CC terminal half of HC, designated as HN), that translocates the therapeutic
 CC agent into the neuronal cell, where the translocation domain is not a HN
 CC domain of a clostridial neurotoxin and is not a fragment or derivative of
 CC a HN domain of a clostridial toxin. Polypeptides of the invention are
 CC useful for the treatment of a disease state associated with neuronal
 CC cells. The polypeptide constructs are useful for delivering therapeutic
 CC substances to neuronal cells. They are useful to treat disorders of the
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
 CC and infection. They are also useful in gene therapy. The present sequence
 CC is C. botulinum C2 enterotoxin translocation domain with tetanus
 CC neurotoxin (TeNT) binding domain used in the exemplification of the
 CC invention.
 CC
 SQ Sequence 1052 AA;

Query Match 97.6%; Score 2397; DB 22; Length 1052;
 Best Local Similarity 99.6%; Pred. No. 2e-185;
 Matches 452; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 SYKNLDGWNEEDIDVILKSTILNDINNDIISDISGNSSVITTPDAQLVPGINGK 69
 DB 597 SYKNLDGWNEEDIDVILKSTILNDINNDIISDISGNSSVITTPDAQLVPGINGK 656
 QY 70 AILVNESSEVIYKRAMDIEYNDMFNFTVFWLRVPKVSASHLEOYGTNEYSIISSMK 129
 DB 657 AILVNESSEVIYKRAMDIEYNDMFNFTVFWLRVPKVSASHLEOYGTNEYSIISSMK 716
 QY 130 KHSLSGGSWSLKGNNLIWTLKDSAGEVROITPDLDPKFNAYLANKWVFITITNDRL 189
 DB 717 KHSLSGGSWSLKGNNLIWTLKDSAGEVROITPDLDPKFNAYLANKWVFITITNDRL 776
 QY 190 SSANLYINGVLMSAEITGLGAIREDNNTTLKLDRCNNNNQYVSIDKFRIEFCALNPKIE 249
 DB 777 SSANLYINGVLMSAEITGLGAIREDNNTTLKLDRCNNNNQYVSIDKFRIEFCALNPKIE 836
 QY 250 EKLKYSYSLITPLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNTTMYLTNAPSTYNGK 309
 DB 837 EKLKYSYSLITPLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNTTMYLTNAPSTYNGK 896
 QY 310 LNIYRRLYNGLKFTIKRTPPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPKGNAFNNIDR 369
 DB 897 LNIYRRLYNGLKFTIKRTPPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPKGNAFNNIDR 956
 QY 370 IRLRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNGQIGNDPNR 429
 DB 957 IRLRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNGQIGNDPNR 1016
 QY 430 DILIASNMYFNHLKDKILGCDWYFVPTDEGWTND 463
 DB 1017 DILIASNMYFNHLKDKILGCDWYFVPTDEGWTND 1050

RESULT 9
 AAE07902
 ID AAE07902 standard; Protein; 1112 AA.
 AC AAE07902;
 DT 01-NOV-2001 (first entry)
 XX
 C. botulinum C2 translocation domain with TeNT binding domain #1.
 DE
 XX Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
 KW tumour; infection; neurodegenerative disease; gene therapy;
 KM botulinum neurotoxin; tetanus neurotoxin; TeNT.
 XX
 OS Clostridium botulinum.
 OS

OS Clostridium tetani.
 XX WO200158936-A2.
 XX 16-AUG-2001.
 XX 04-DEC-2000; 2000MO-GB04644.
 XX 02-DEC-1999; 99GB-0028530.
 PR 07-APR-2000; 2000GB-0008658.
 XX
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
 PI Shone CC, Sutton JM, Silman N;
 WIPI: 2001-514643/56.
 XX
 PT New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -
 PS
 XX Example 2; Page 49; 50pp; English.

The invention relates to a non toxic polypeptide, for delivery of a
 CC therapeutic agent to a neuronal cell, which comprises a binding domain
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
 CC as Hc) that binds to the neuronal cell and a translocation domain (amino
 CC terminal half of HC, designated as HN), that translocates the therapeutic
 CC agent into the neuronal cell, where the translocation domain is not a HN
 CC a HN domain of a clostridial neurotoxin and is not a fragment or derivative of
 CC a useful for the treatment of a disease state associated with neuronal
 CC cells. The polypeptide constructs are useful for delivering therapeutic
 CC substances to neuronal cells. They are useful for treating disorders of the
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
 CC and infection. They are also useful in gene therapy. The present sequence
 CC is C. botulinum C2 enterotoxin translocation domain with tetanus
 CC neurotoxin (Tent) binding domain used in the exemplification of the
 CC invention.

XX Sequence 1112 AA;

Query Match 97.6%; Score 2397; DB 22; Length 1112;
 Best Local Similarity 99.6%; Pred. No. 2.2e-185;
 Matches 451; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 11 YSKNDQWVNEEDIDVILKSTIILNDINNDIISDGSFSSVITPPDQVPGINGKA 70
 DB 658 FTKNDQWVNEEDIDVILKSTIILNDIISDGSFSSVITPPDQVPGINGKA 717
 QY 71 IHLVNNSESEVIVHKAMDIEYNDMFNFTVFWLRVPKVSASHLEQGTNEYSIISSMKK 130
 DB 718 IHLVNNSESEVIVHKAMDIEYNDMFNFTVFWLRVPKVSASHLEQGTNEYSIISSMKK 777
 QY 131 HSLISGSGSVSLKGNLITWLKDSAGEVROITFRDLPKFNAYLANKKWFTITNDRLS 190
 DB 778 HSLISGSGSVSLKGNLITWLKDSAGEVROITFRDLPKFNAYLANKKWFTITNDRLS 837
 QY 191 SANLYINGVLMGSAEITGGAIRREDNNITLKDRCNNNNQOYVSIIDKFRIFCKALNPKREIE 250
 DB 838 SANLYINGVLMGSAEITGGAIRREDNNITLKDRCNNNNQOYVSIIDKFRIFCKALNPKREIE 897
 QY 251 KLYTSTYLSTFLRDWGNPLRDYTEYLLIPVASSSKDVOLEKNTIDMYLTNAPSYNGKL 310
 DB 898 KLYTSTYLSTFLRDWGNPLRDYTEYLLIPVASSSKDVOLEKNTIDMYLTNAPSYNGKL 957
 QY 311 NIYYRLNGLKFLIKRTYPPNNEIDSFYKSGDFIKLYSYNNNEHTVGPCKGNANNNND 370
 DB 958 NIYYRLNGLKFLIKRTYPPNNEIDSFYKSGDFIKLYSYNNNEHTVGPCKGNANNNND 1017
 QY 371 RILRVGNAPGIPLVKKMEAVKLRDLKTYVOLKLYLDKKNASGLGIVTNGOIGNPENR 430
 DB 1018 RILRVGNAPGIPLVKKMEAVKLRDLKTYVOLKLYLDKKNASGLGIVTNGOIGNPENR 1077

QY 431 ILIASMWYFNHLKDKITLGCWMYFVPTDEGWTND 463
 DB 1078 ILIASMWYFNHLKDKITLGCWMYFVPTDEGWTND 1110

RESULT 10
 AAR12471
 ID AAR12471 standard; Protein; 452 AA.
 XX
 AC AAR12471;
 XX
 DT 05-AUG-1991 (first entry)
 XX
 DE Tetanus toxin fragment C encoded by gene with increased G+C content.
 XX
 KM Terminator; vaccine.
 XX
 OS Synthetic.
 XX
 PN EP430645-A.
 XX
 PD 05-JUN-1991.
 XX
 PE 27-NOV-1990; 90EP-0312870.
 XX
 PR 17-MAR-1990; 90GB-0006097.
 PR 28-NOV-1989; 89GB-0026832.
 XX
 PA (WELL) WELLCOME FOUNDATION LTD.
 XX
 PI Makoff AJ, Romanos MA, Clare JJ, Fairweather NF;
 XX
 DR WPI: 1991-166115/23.
 DR N-PSDB; AAO12121.

PT DNA sequence encoding tetanus toxin fragment C - useful in the
 PT manufacture of vaccines for immunity to tetanus utilising yeast
 PT as host organism.
 XX
 PS Disclosure; Fig 2; 50pp; English.

XX The (G+C) content of the synthetic gene is increased by 47% wrt the
 CC native sequence. This eliminates six "terminator" regions which
 CC were found to be present in (A+T) rich regions. The terminators
 CC (termination/endo-nucleolytic processing/polyadenylation sites)
 CC were previously responsible for incomplete transcription of the
 CC mRNA. The elimination of these elements (using codon degeneracy)
 CC provided for successful expression in yeast of the tetanus toxin
 CC fragment C.

XX Sequence 452 AA;

Query Match 97.4%; Score 2393; DB 12; Length 452;
 Best Local Similarity 100.0%; Pred. No. 1.2e-185;
 Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 KNDQWVNEEDIDVILKSTIILNDINNDIISDGSFSSVITPPDQVPGINGKA 72
 DB 2 KNDQWVNEEDIDVILKSTIILNDINNDIISDGSFSSVITPPDQVPGINGKA 61
 QY 73 LVNNESEVIVHKAMDIEYNDMFNFTVFWLRVPKVSASHLEQGTNEYSIISSMKK 132
 DB 62 LVNNESEVIVHKAMDIEYNDMFNFTVFWLRVPKVSASHLEQGTNEYSIISSMKK 121
 QY 133 LSTGSGSVSLKGNLITWLKDSAGEVROITFRDLPKFNAYLANKKWFTITNDRLSSA 192
 DB 122 LSTGSGSVSLKGNLITWLKDSAGEVROITFRDLPKFNAYLANKKWFTITNDRLSSA 181
 QY 193 NLVINGVLMGSAEITGGAIRREDNNITLKDRCNNNNQOYVSIIDKFRIFCKALNPKREIE 252
 DB 182 NLVINGVLMGSAEITGGAIRREDNNITLKDRCNNNNQOYVSIIDKFRIFCKALNPKREIE 241

```

OY 253 YTSYLSITFLRDFMGNPLRYDTEYYLIIPVASSSKDYOLKNTIDYMYLTNAPSYTNGKLN 312
DB 242 YTSYLSITFLRDFMGNPLRYDTEYYLIIPVASSSKDYOLKNTIDYMYLTNAPSYTNGKLN 301
OY 313 YRRLYLNGIKFKIRKRYTPNNEIDSEVKSDFIKLYSYNNNEHIVGYPKDGNAFNMLDRI 372
DB 302 YRRLYLNGIKFKIRKRYTPNNEIDSEVKSDFIKLYSYNNNEHIVGYPKDGNAFNMLDRI 361
OY 373 LRVGNAGGIPLYKRMKAVKLRDLKTSYVOLKLYDPKNAISLGLVGHNGOIGNDPNDRI 432
DB 362 LRVGNAGGIPLYKRMKAVKLRDLKTSYVOLKLYDPKNAISLGLVGHNGOIGNDPNDRI 421
OY 433 IASNWFNHLKDKILGCDWYFVPTDEGWTND 463
DB 422 IASNWFNHLKDKILGCDWYFVPTDEGWTND 452

RESULT 11
AAB31429
ID AAB31429 standard; protein: 661 AA.
XX
AC AAB31429;
XX
DT 26-APR-2001 (first entry)
XX
DE Shed antigen-specific B cell antigen linked to tetanus toxin fragment C.
XX
KW Vaccine; shed antigen-specific B cell; idiotypic antibody.
XX
KW Immune complex-mediated disease; autoimmune disease; tetanus protein;
XX
KW humoral immune response; cancer.
XX
OS Synthetic.
XX
OS Clostridium tetani.
XX
PN WO200076319-A1.
XX
PD 21-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16677.
XX
PR 16-JUN-1999; 1999US-0139521.
XX
PR 15-JUN-2000; 2000US-0594985.
XX
PA (BIOC-) BIOCRYSTAL LTD.
XX
PI Nelson MB, Barbera-Guillem E;
XX
WPI: 2001-080635/09.
XX
DR
XX
PT Inducing an immune response against shed antigen-specific B cell
PT polynucleotides encoding an idiotypic determinant or peptides
PT comprising an idiotypic determinant

Example 2; Page 73-76; 81pp; English.
XX
PS The present sequence represents a fusion protein, comprising a protein
XX used for immunising against shed antigen-specific B cells linked to a
XX fragment of tetanus protein. It is used in vaccines of the invention. The
XX specification describes a method for inducing an immune response reactive
XX with idiotypes present on shed antigen-specific B cells (SAb) of an
XX individual. The method involves administering a vaccine formulation
XX comprising polynucleotide encoding an idiotype of an antibody that binds
XX to an epitope of shed antigen. The method is useful for inducing an
XX immune response reactive with idiotypes present on SAb of an individual.
XX The method is useful for depleting shed antigen-specific B cells, and
XX for treating immune complex-mediated disease progression in organ
XX specific autoimmune disease exacerbated by humoral immune response
XX against groups expressed on shed antigen, or by plasma cell production
XX of antibodies against groups of shed antigen. It is useful in cancer
XX therapy and for treating autoimmune disease.
XX
SQ Sequence 661 AA;

```

```

Query Match 97.2%; Score 2387.5; DB 22; Length 661;
Best Local Similarity 95.2%; Pred. No. 6e-185;
Matches 453; Conservative 2; Mismatches 2; Indels 19; Gaps 1;

OY 7 IPEYS-----KNIDCVNDEEDIDVILKSTIILNDINDIISDI 47
DB 186 VPFRRSGSGVWDLFLSHNPNGGLEKNIDCVNDEEDIDVILKSTIILNDINDIISDI 245
OY 48 SGFNSSVITTPDAQVGLNGINKALHIVNSESSEVYVHKAMDIEYDMNNTFVSFWLRYP 107
DB 246 SGFNSSVITTPDAQVGLNGINKALHIVNSESSEVYVHKAMDIEYDMNNTFVSFWLRYP 305
OY 108 KVSASHLEOYGTNNEYSLISSMKHSLSIGSGWSYSLKGNLITWLKDSAGEVROITFRDL 167
DB 306 KVSASHLEOYGTNNEYSLISSMKHSLSIGSGWSYSLKGNLITWLKDSAGEVROITFRDL 365
OY 168 PDKFNAYLANKWVFITTTNDRILSSANLYINGVLMGSAEITGIGAIREDNNITLKDRCNN 227
DB 366 PDKFNAYLANKWVFITTTNDRILSSANLYINGVLMGSAEITGIGAIREDNNITLKDRCNN 425
OY 228 NNOYVSIDKFRIFCKALNPKETIEKLYSTLSTFLRDFMGNPLRYDTEYYLIIPVASSSKD 287
DB 426 NNOYVSIDKFRIFCKALNPKETIEKLYSTLSTFLRDFMGNPLRYDTEYYLIIPVASSSKD 485
OY 288 VOLKNTIDYMYLTNAPSYTNGKLNLYRRLYNGIKFKIRKRYTPNNEIDSEVKSDFIKLY 347
DB 486 VOLKNTIDYMYLTNAPSYTNGKLNLYRRLYNGIKFKIRKRYTPNNEIDSEVKSDFIKLY 545
OY 348 VSYNNNEHIVGYPKDGNAFNMLDRI LRVGNAGGIPLYKRMKAVKLRDLKTSYVOLKLYD 407
DB 546 VSYNNNEHIVGYPKDGNAFNMLDRI LRVGNAGGIPLYKRMKAVKLRDLKTSYVOLKLYD 605
OY 408 DKNASLGLVGHNGOIGNDPNDRI IASNWFNHLKDKILGCDWYFVPTDEGWTND 463
DB 606 DKNASLGLVGHNGOIGNDPNDRI IASNWFNHLKDKILGCDWYFVPTDEGWTND 661

RESULT 12
AAB31427
ID AAB31427 standard; protein: 453 AA.
XX
AC AAB31427;
XX
DT 20-APR-2001 (first entry)
XX
DE Amino acid sequence of tetanus toxin fragment C.
XX
KW Vaccine; shed antigen-specific B cell; idiotypic antibody.
XX
KW Immune complex-mediated disease; autoimmune disease; tetanus protein;
XX
KW humoral immune response; cancer.
XX
OS Clostridium tetani.
XX
PN WO200076319-A1.
XX
PD 21-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16677.
XX
PR 16-JUN-1999; 99US-0139521.
XX
PR 15-JUN-2000; 2000US-0594985.
XX
PA (BIOC-) BIOCRYSTAL LTD.
XX
PI Nelson MB, Barbera-Guillem E;
XX
WPI: 2001-080635/09.
XX
DR
XX
PT Inducing an immune response against shed antigen-specific B cell
PT polynucleotides encoding an idiotypic determinant or peptides
PT comprising an idiotypic determinant

```

```
XX Example 2; Page 72-73; 81pp; English.
PS
XX
CC The present sequence represents a fragment of tetanus protein, which
CC is used as an immunostimulatory protein in vaccines of the invention. The
CC specification describes a method for inducing an immune response reactive
CC with idiotypes present on shed antigen-specific B cells (SAB) of an
CC individual. The method involves administering a vaccine formulation
CC comprising polynucleotide encoding an idio type of an antibody that binds
CC to an epitope of shed antigen. The method is useful for inducing an
CC immune response reactive with idiotypes present on SAB of an individual.
CC The method is useful for depleting shed antigen-specific B cells, and
CC for treating immune complex-mediated disease progression in organ
CC specific autoimmune disease exacerbated by humoral immune response
CC against groups expressed on shed antigen, or by plasma cell production
CC of antibodies against groups of shed antigen. It is useful in cancer
CC therapy and for treating autoimmune disease.
XX
XX Sequence 453 AA:
Query Match 97.1%; Score 2385; DB 22; Length 453;
Best Local Similarity 99.8%; Pred. No. 5.5e-185;
Matches 450; Conservative 0; Mismatches 1; Indels 0; Gaps 0:
QY 13 KNDCWVDNEEDIDVILKKSITLNDINNDISISGNSSVITPPAQLYVGINGKAH 72
DB 3 KNDCWVDNEEDIDVILKKSITLNDINNDISISGNSSVITPPAQLYVGINGKAH 62
QY 73 LVNSESSEVIVHKAMDIEYNDMENNFTVSEFWLRVPKVSASHLEQGTNEYSTISSMKKHS 132
DB 63 LVNSESSEVIVHKAMDIEYNDMENNFTVSEFWLRVPKVSASHLEQGTNEYSTISSMKKHS 122
QY 133 LSTGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPKFNAYLANKWVFTITNDRLSSA 192
DB 123 LSTGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPKFNAYLANKWVFTITNDRLSSA 182
QY 193 NLYTNGVLMGSAEITGIGALREDNNITLKLDRGNNNNOYVSIDKFRIFCKALNPEIEKLYT 252
DB 183 NLYTNGVLMGSAEITGIGALREDNNITLKLDRGNNNNOYVSIDKFRIFCKALNPEIEKLYT 242
QY 253 YTSYSLITFLRDFWGNPLRYDTEYLLIPVASSSKDVOUKNTDWMYLTNAPSYNGKLN 312
DB 243 YTSYSLITFLRDFWGNPLRYDTEYLLIPVASSSKDVOUKNTDWMYLTNAPSYNGKLN 302
QY 313 YTRRLYNGLKFTIKRTPPNNEIDSEVKSQDFIKLYSYNNNEHIVGYPKDGNAPNNDRI 372
DB 303 YTRRLYNGLKFTIKRTPPNNEIDSEVKSQDFIKLYSYNNNEHIVGYPKDGNAPNNDRI 362
QY 373 LRVGYNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGIVGTHNGOIGNDPNRDL 432
DB 363 LRVGYNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGIVGTHNGOIGNDPNRDL 422
QY 433 IASNMYFNHLKDKITLGCDDWYFVPTDEGWTND 463
DB 423 IASNMYFNHLKDKITLGCDDWYFVPTDEGWTND 453
RESULT 13
AAE07897
ID AAE07897 standard; Protein; 605 AA.
XX
AC AAE07897;
XX
DT 01-NOV-2001 (first entry)
XX
DE Modified clostridial heavy chain fragment #4.
XX
KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
KW tumour; infection; neurodegenerative disease; gene therapy; chimeric;
KW diphtheria neurotoxin; tetanus neurotoxin; TeNT.
XX
OS Chimeric - Corynebacterium diphtheriae.
OS Chimeric - Clostridium tetani.
```

```
XX
PN WO200158936-A2.
XX
PD 16-AUG-2001.
XX
PF 04-DEC-2000; 2000WO-GB04644.
XX
PR 02-DEC-1999; 99GB-0028530.
PR 07-APR-2000; 2000GB-0008658.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
PI Shone CC, Sutton JM, Silman N;
XX
DR WPI, 2001-514643/56.
XX
PT New non toxic polypeptide for delivery of a therapeutic agent for the
PT treatment of a CNS disorder comprising a binding domain that
PT translocates the therapeutic agent into the neuronal cells -
XX
PS Example 2; Page 45; 50pp; English.
XX
CC The invention relates to a non toxic polypeptide, for delivery of a
CC therapeutic agent to a neuronal cell, which comprises a binding domain
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC as HC) that binds to the neuronal cell and a translocation domain (amino
CC terminal half of HC, designated as HN), that translocates the therapeutic
CC agent into the neuronal cell, where the translocation domain is not a HN
CC domain of a clostridial neurotoxin and is not a fragment or derivative of
CC a HN domain of a clostridial toxin. Polypeptides of the invention are
CC useful for the treatment of a disease state associated with neuronal
CC cells. The polypeptide constructs are useful for delivering therapeutic
CC CNS instances to neuronal cells. They are useful to treat disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is modified clostridial heavy chain fragment. This sequence is
CC constructed by fusing the truncated binding domain of tetanus neurotoxin
CC (TeNT) with translocation domain of diphtheria neurotoxin.
XX
XX Sequence 605 AA:
Query Match 84.4%; Score 2073; DB 22; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.7e-159;
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0:
QY 75 NNESESEVIVHKAMDIEYNDMENNFTVSEFWLRVPKVSASHLEQGTNEYSTISSMKKHSLS 134
DB 215 NNESESEVIVHKAMDIEYNDMENNFTVSEFWLRVPKVSASHLEQGTNEYSTISSMKKHSLS 124
QY 135 IGSQWSVSLKGNLIMTLKDSAGEVROITFRDLPKFNAYLANKWVFTITNDRLSSANL 194
DB 275 IGSQWSVSLKGNLIMTLKDSAGEVROITFRDLPKFNAYLANKWVFTITNDRLSSANL 184
QY 195 YINGVLMGSAEITGIGALREDNNITLKLDRGNNNNOYVSIDKFRIFCKALNPEIEKLYT 254
DB 335 YINGVLMGSAEITGIGALREDNNITLKLDRGNNNNOYVSIDKFRIFCKALNPEIEKLYT 244
QY 255 SYLSTITFLRDFWGNPLRYDTEYLLIPVASSSKDVOUKNTDWMYLTNAPSYNGKLN 314
DB 395 SYLSTITFLRDFWGNPLRYDTEYLLIPVASSSKDVOUKNTDWMYLTNAPSYNGKLN 304
QY 315 RRLYNGLKFTIKRTPPNNEIDSEVKSQDFIKLYSYNNNEHIVGYPKDGNAPNNDRI 374
DB 455 RRLYNGLKFTIKRTPPNNEIDSEVKSQDFIKLYSYNNNEHIVGYPKDGNAPNNDRI 364
QY 375 VGYNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGIVGTHNGOIGNDPNRDL 434
DB 515 VGYNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGIVGTHNGOIGNDPNRDL 424
QY 435 SNMYFNHLKDKITLGCDDWYFVPTDEGWTND 463
DB 575 SNMYFNHLKDKITLGCDDWYFVPTDEGWTND 603
```

[illegible]

Qy	148	LWTLKDSAGEVROITTR-----DLPRKFNNLYLANKKAVFTITINDRLSSANKIYNGVMG	202
Dd	344	IWTLTDDTAGNKKOVLFEWYTMOMISD---YI-NKMIPTVITNNRGRNSIRIYNGILMD	398
Qy	203	SAETIGGATIEDNNITLKLDRCNNNOYASIDKFRJECALNPKIELEKYTSYLSITTEL	262
Dd	399	EKSIISNLEDIHVSODMLFKYIGVC-NDTRIVSIDKFRJFCALNPKIELEKYTSYLSITTEL	457
Qy	263	RDFWGNPLRYDEYLLIFVASSSKRDVOLKNTIDMYLTLNAPSYYNGKLNITYRRLYNGLK	322
Dd	458	RDFWGNPLRYDEYLLIFVASSSKRDVOLKNTIDMYLTLNAPSYYNGKLNITYRRLYNGLK	517
Qy	323	FLIKRYTPENNEIDSEVKSDFETKLYVSYNNEHIVGTBPKGNMAFNMLDRLLRGSYNAPI	382
Dd	518	FLIKRYTPENNEIDSEVKSDFETKLYVSYNNEHIVGTBPKGNMAFNMLDRLLRGSYNAPI	577
Qy	383	PLYKMEAVKLRLDKTYSVQKLKDDKASGLYGTINNGOQGNPNRDLILASWNYFNHL	442
Dd	578	PLYKMEAVKLRLDKTYSVQKLKDDKASGLYGTINNGOQGNPNRDLILASWNYFNHL	637
Qy	443	KDKILIGDMYFVPPDEGNTND	463
Dd	638	KDKILIGDMYFVPTDEGNTND	658

RESULT 15
AAE07896
AAE07896 standard; Protein; 448 AA.
AAE07896;
01-NOV-2001 (first entry)
Modified clostridial heavy chain fragment #3.
Neuronal cell; binding domain; translocation domain; stroke; epilepsy; tumor; infection; neurodegenerative disease; gene therapy; chimeric; diphtheria neurotoxin; tetanus neurotoxin; TeNT.
Chimeric - Corynebacterium diphtheriae.
Chimeric - Clostridium tetani.
MO200158936-A2.
16-AUG-2001.
04-DEC-2000; 2000MO-GB04644.
02-DEC-1999; 99GB-0028530.
07-APR-2000; 2000GB-0008658.
(MIGR-) MICROBIOLOGICAL RES AUTHORITY.
Shone CC, Sutton JM, Silman N;
WPI; 2001-514643/56.
New non toxic polypeptide for delivery of a therapeutic agent for the treatment of a CNS disorder comprising a binding domain that translocates the therapeutic agent into the neuronal cells -
Example 2; Page 45; 50pp; English.
The invention relates to a non toxic polypeptide, for delivery of a therapeutic agent to a neuronal cell, which comprises a binding domain (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated as Hc) that binds to the neuronal cell and a translocation domain (amino terminal half of HC, designated as HN), that translocates the therapeutic agent into the neuronal cell, where the translocation domain is not a HN domain of a clostridial neurotoxin and is not a fragment or derivative of a HN domain of a clostridial toxin. Polypeptides of the invention are useful for the treatment of a disease state associated with neuronal cells. The polypeptide constructs are useful for delivering therapeutic

CC substances to neuronal cells. They are useful to treat disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is modified clostridial heavy chain fragment. This sequence is
CC constructed by fusing the binding domain II of tetanus neurotoxin (TnT)
CC with translocation domain of diphtheria neurotoxin.
XX

SQ Sequence 448 AA;

Query Match 51.3%; Score 1261; DB 22; Length 448;

Best Local Similarity 100.0%; Pred. No. 8e-94;

Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	231	YVSIDKFRIFCKALNPKIEKLYTSYSTFLRDFWGNPLRYDTEYLLIPVASSSKDVQL	290
DB	214	YVSIDKFRIFCKALNPKIEKLYTSYSTFLRDFWGNPLRYDTEYLLIPVASSSKDVQL	273
QY	291	KNITDWMYLTNAPSYTNGKLNITYRRRLYNGLKFTIKRTPNNEIDSFVKSQDFTKLYVS	350
DB	274	KNITDWMYLTNAPSYTNGKLNITYRRRLYNGLKFTIKRTPNNEIDSFVKSQDFTKLYVS	333
QY	351	NNNEHIVGYPKDGNAFNNLDRIILRVGNAPGIPLYKKMEAVKLRDLKTYVOLKLYDDKN	410
DB	334	NNNEHIVGYPKDGNAFNNLDRIILRVGNAPGIPLYKKMEAVKLRDLKTYVOLKLYDDKN	393
QY	411	ASIGLVGTHNGQIGNDPNRDILLASNMYFNHLKDKILGCDWYFVPTDEGWTND	463
DB	394	ASIGLVGTHNGQIGNDPNRDILLASNMYFNHLKDKILGCDWYFVPTDEGWTND	446

Search completed: December 1, 2002, 11:21:13
Job time : 70 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2002, 11:20:37 ; Search time 26 Seconds

(Without alignments)
523.954 Million cell updates/sec

Title: US-09-816-467-2

Perfect score: 2456
Sequence: 1 MFESTPIPFYSKNDICWD.....DKILGDMYFVPTDEGTND 463

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2.6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2.6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2.6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2.6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2.6/ptodata/1/1aa/PTCUS.COMB.pep:*
6: /cgn2.6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query length	ID	Description
1	2451	99.8	853 4 US-08-913-880C-17	Sequence 17, Appl
2	2451	99.8	858 4 US-08-913-880C-16	Sequence 16, Appl
3	2451	99.8	860 4 US-08-913-880C-15	Sequence 15, Appl
4	2451	99.8	862 4 US-08-913-880C-14	Sequence 14, Appl
5	2451	99.8	865 4 US-08-913-880C-13	Sequence 13, Appl
6	2451	99.8	866 4 US-08-913-880C-12	Sequence 12, Appl
7	2451	99.8	874 4 US-08-913-880C-11	Sequence 11, Appl
8	2451	99.8	875 4 US-08-913-880C-10	Sequence 10, Appl
9	2451	99.8	1315 4 US-08-913-880C-1	Sequence 5, Appl
10	2440	99.3	618 1 US-07-618-312A-5	Sequence 4, Appl
11	2393	97.4	452 1 US-07-618-312A-4	Sequence 2, Appl
12	2380	96.9	452 1 US-07-618-312A-2	Sequence 2, Appl
13	2380	96.9	452 1 US-07-618-312A-2	Sequence 2, Appl
14	2380	96.9	452 1 US-08-280-228-2	Sequence 8, Appl
15	2345	95.5	452 1 US-08-110-786A-8	Sequence 28, Appl
16	653.5	26.6	1296 2 US-08-405-496A-28	Sequence 28, Appl
17	653.5	26.6	1296 2 US-08-405-496A-28	Sequence 28, Appl
18	653.5	26.6	1296 2 US-08-405-496A-28	Sequence 28, Appl
19	627.5	25.5	438 1 US-08-480-604A-23	Sequence 23, Appl
20	627.5	25.5	438 1 US-08-405-496A-23	Sequence 23, Appl
21	627.5	25.5	438 1 US-08-405-496A-23	Sequence 23, Appl
22	627.5	25.5	438 1 US-08-405-496A-23	Sequence 23, Appl
23	627.5	25.5	462 2 US-08-480-604A-26	Sequence 26, Appl
24	627.5	25.5	462 2 US-08-405-496A-26	Sequence 26, Appl
25	478	19.5	1169 4 US-08-915-136-26	Sequence 20, Appl
26	220	9.0	44 1 US-08-110-786A-5	Sequence 5, Appl
27	168	6.8	31 5 PCT-US93-11703-64	Sequence 64, Appl

28	163	6.6	140 4 US-08-446-114A-22	Sequence 22, Appl
29	127	5.2	2391 2 US-08-446-855A-2	Sequence 2, Appl
30	127	5.2	2391 4 US-09-150-741-2	Sequence 2, Appl
31	125	5.1	20 1 US-08-231-437-1	Sequence 1, Appl
32	125	5.1	20 5 PCT-US93-06175-1	Sequence 1, Appl
33	125	5.1	21 1 PCT-US93-06175-1	Sequence 2, Appl
34	125	5.1	21 1 PCT-US93-06175-2	Sequence 2, Appl
35	119.5	4.9	2366 1 US-08-480-604A-10	Sequence 10, Appl
36	119.5	4.9	2366 2 US-08-405-496A-10	Sequence 10, Appl
37	119.5	4.9	2366 4 US-08-915-136-10	Sequence 10, Appl
38	119.5	4.9	2366 4 US-08-957-310-10	Sequence 10, Appl
39	115	4.7	993 4 US-08-836-687B-30	Sequence 30, Appl
40	114	4.6	32 1 US-08-446-692-14	Sequence 14, Appl
41	114	4.6	32 2 US-08-488-351A-14	Sequence 14, Appl
42	112.5	4.6	1005 4 US-09-206-942-41	Sequence 41, Appl
43	112.5	4.6	1011 4 US-09-206-942-39	Sequence 39, Appl
44	112	4.6	21 1 US-07-610-525-1	Sequence 1, Appl
45	112	4.6	21 2 US-08-661-052-12	Sequence 12, Appl

ALIGNMENTS

```
RESULT 1
US-08-913-880C-17
; Sequence 17, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihito
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 17
; LENGTH: 853
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 463 to 1315 of SEQ ID NO: 1
US-08-913-880C-17

Query Match          99.8%: Score 2451; DB 4; Length 853;
Best Local Similarity 100.0%: Pred. No. 2.7e-189;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VFSTPIPFYSKNDICWDNEEDIDVILKSTIINDIISDGFNSVITYPPDAQ 61
|||||
DB 392 VFSTPIPFYSKNDICWDNEEDIDVILKSTIINDIISDGFNSVITYPPDAQ 451
|||||
QY 62 LVPGINKATIHIVNSESSEVYVHKAMDEYDMRNFTVSWLKPYSASHLEQGTNE 121
|||||
DB 452 LVPGINKATIHIVNSESSEVYVHKAMDEYDMRNFTVSWLKPYSASHLEQGTNE 511
|||||
QY 122 YSIISMKKHSLSIGSGWSYSLKGNLITWLKDSAGEVROITFRDLDPKFNAYLANKWYF 181
|||||
DB 512 YSIISMKKHSLSIGSGWSYSLKGNLITWLKDSAGEVROITFRDLDPKFNAYLANKWYF 571
|||||
QY 182 ITTNRRLSANLIYINGVLMGSAEITGLGAIREDNNITLTKIDRCNNNNQVYSIDKFRIFC 241
|||||
DB 572 ITTNRRLSANLIYINGVLMGSAEITGLGAIREDNNITLTKIDRCNNNNQVYSIDKFRIFC 631
|||||
QY 242 KALNPKIEIKLYSYSTIFLRDFGNGPLRDYDEYLIIPVASSSKDVOLKITTYMYLTN 301
|||||
DB 632 KALNPKIEIKLYSYSTIFLRDFGNGPLRDYDEYLIIPVASSSKDVOLKITTYMYLTN 691
|||||
QY 302 ABSYTNGLNTIYRRLYNGIKFKITKRYTPNNEIDSFPKSGDFIKLYSYNNNEHIVGYPK 361
|||||
DB 692 ABSYTNGLNTIYRRLYNGIKFKITKRYTPNNEIDSFPKSGDFIKLYSYNNNEHIVGYPK 751
|||||
QY 362 DGNAFNNIDRIILRVGNAGCIPLXKKWEAKVLDKITYSVOLKLYDKNASLGLVGHNG 421
|||||
```

```

Db 752 DGNMFNNIDRLRVGNAPGIPLYKKMEAVKLRDLKTSVOLKLYDXKASLGIVGTHNG 811
OY 422 QIGNDPNRDLIASNMWYFNHLKDKILGCDWYFVPTDEGWTND 463
Db 812 QIGNDPNRDLIASNMWYFNHLKDKILGCDWYFVPTDEGWTND 853

RESULT 2
US-08-913-880C-16
; Sequence 16, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiko
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 16
; LENGTH: 858
; TYPE: PRP
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 458 to 1315 of SEQ ID NO: 1
US-08-913-880C-16

Query Match          99.8%; Score 2451; DB 4; Length 858;
Best Local Similarity 100.0%; Pred. No. 2,7e-189;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VFSTPIPFYSKNDLDCWVDNEEDIDVILKSTILNDINNDIISDSIGFNSSVITTPDAQ 61
Db 397 VFSTPIPFYSKNDLDCWVDNEEDIDVILKSTILNDINNDIISDSIGFNSSVITTPDAQ 456
OY 62 LVPGINGKAIHLVNNSESEYIVHKAMDIENDMFNFTVSFWLRVPKVSASHLEOYGTNE 121
Db 457 LVPGINGKAIHLVNNSESEYIVHKAMDIENDMFNFTVSFWLRVPKVSASHLEOYGTNE 516
OY 122 YSISSMKRHSLSIGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPRKFNAYLANKWVF 181
Db 517 YSISSMKRHSLSIGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPRKFNAYLANKWVF 576
OY 182 ITTNDRLSSANLYINGVLMSAEITGLAIRDNNITTLKDRCNNNNNOYVSIDKFRIFC 241
Db 577 ITTNDRLSSANLYINGVLMSAEITGLAIRDNNITTLKDRCNNNNNOYVSIDKFRIFC 636
OY 242 KALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYVLIIPVASSSKDVOLKNITDYMTLN 301
Db 637 KALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYVLIIPVASSSKDVOLKNITDYMTLN 696
OY 302 APSYTNKLNITRYRLYNGLKFTIKRTYPNNEIDSFVKSDFIKLYSYNNNEHIVGYPK 361
Db 697 APSYTNKLNITRYRLYNGLKFTIKRTYPNNEIDSFVKSDFIKLYSYNNNEHIVGYPK 756
OY 362 DGNMFNNIDRLRVGNAPGIPLYKKMEAVKLRDLKTSVOLKLYDXKASLGIVGTHNG 421
Db 757 DGNMFNNIDRLRVGNAPGIPLYKKMEAVKLRDLKTSVOLKLYDXKASLGIVGTHNG 816
OY 422 QIGNDPNRDLIASNMWYFNHLKDKILGCDWYFVPTDEGWTND 463
Db 817 QIGNDPNRDLIASNMWYFNHLKDKILGCDWYFVPTDEGWTND 858

RESULT 3
US-08-913-880C-15
; Sequence 15, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiko
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P

```

```

; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 15
; LENGTH: 860
; TYPE: PRP
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 456 to 1315 of SEQ ID NO: 1
US-08-913-880C-15

Query Match          99.8%; Score 2451; DB 4; Length 860;
Best Local Similarity 100.0%; Pred. No. 2,7e-189;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VFSTPIPFYSKNDLDCWVDNEEDIDVILKSTILNDINNDIISDSIGFNSSVITTPDAQ 61
Db 397 VFSTPIPFYSKNDLDCWVDNEEDIDVILKSTILNDINNDIISDSIGFNSSVITTPDAQ 458
OY 62 LVPGINGKAIHLVNNSESEYIVHKAMDIENDMFNFTVSFWLRVPKVSASHLEOYGTNE 121
Db 457 LVPGINGKAIHLVNNSESEYIVHKAMDIENDMFNFTVSFWLRVPKVSASHLEOYGTNE 518
OY 122 YSISSMKRHSLSIGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPRKFNAYLANKWVF 181
Db 519 YSISSMKRHSLSIGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPRKFNAYLANKWVF 578
OY 182 ITTNDRLSSANLYINGVLMSAEITGLAIRDNNITTLKDRCNNNNNOYVSIDKFRIFC 241
Db 579 ITTNDRLSSANLYINGVLMSAEITGLAIRDNNITTLKDRCNNNNNOYVSIDKFRIFC 638
OY 242 KALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYVLIIPVASSSKDVOLKNITDYMTLN 301
Db 639 KALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYVLIIPVASSSKDVOLKNITDYMTLN 698
OY 302 APSYTNKLNITRYRLYNGLKFTIKRTYPNNEIDSFVKSDFIKLYSYNNNEHIVGYPK 361
Db 699 APSYTNKLNITRYRLYNGLKFTIKRTYPNNEIDSFVKSDFIKLYSYNNNEHIVGYPK 758
OY 362 DGNMFNNIDRLRVGNAPGIPLYKKMEAVKLRDLKTSVOLKLYDXKASLGIVGTHNG 421
Db 759 DGNMFNNIDRLRVGNAPGIPLYKKMEAVKLRDLKTSVOLKLYDXKASLGIVGTHNG 818
OY 422 QIGNDPNRDLIASNMWYFNHLKDKILGCDWYFVPTDEGWTND 463
Db 819 QIGNDPNRDLIASNMWYFNHLKDKILGCDWYFVPTDEGWTND 860

RESULT 4
US-08-913-880C-14
; Sequence 14, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiko
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 14
; LENGTH: 862
; TYPE: PRP
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 454 to 1315 of SEQ ID NO: 1
US-08-913-880C-14

Query Match          99.8%; Score 2451; DB 4; Length 862;
Best Local Similarity 100.0%; Pred. No. 2,7e-189;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VFSTPIPFYSKNDLDCWVDNEEDIDVILKSTILNDINNDIISDSIGFNSSVITTPDAQ 61

```

```

Db 401 VFSPIPFYSKNDLCWVNEEDIDVILKSTILNDINNDIISDGSFNSVITYPPAQ 460
QY 62 LVPGINGKAHLVNNESSEVIYHKAMDEIYNDMPNFVSWLRPKVSASHLEOYGTNE 121
Db 461 LVPGINGKAHLVNNESSEVIYHKAMDEIYNDMPNFVSWLRPKVSASHLEOYGTNE 520
QY 122 YSISSMKKHSLSIGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPDKFNAYLANKWVF 181
Db 521 YSISSMKKHSLSIGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPDKFNAYLANKWVF 580
QY 182 ITTNDRLSSANLYINGVLGMSAETTGAIREDNNITLKDRCNNNNQYVSIDKFRIFC 241
Db 581 ITTNDRLSSANLYINGVLGMSAETTGAIREDNNITLKDRCNNNNQYVSIDKFRIFC 640
QY 242 KALNPKEIEKLYTSITFLRDFWGNPLRYDTEYLLPVASSSKDVOLKNTIDVWYLTN 301
Db 641 KALNPKEIEKLYTSITFLRDFWGNPLRYDTEYLLPVASSSKDVOLKNTIDVWYLTN 700
QY 302 APSYTNGLKLNYYRRLYNGLKFTIKRYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 361
Db 701 APSYTNGLKLNYYRRLYNGLKFTIKRYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 760
QY 362 DGNAFNNLDRILRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 421
Db 761 DGNAFNNLDRILRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 820
QY 422 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 463
Db 821 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 862

```

RESULT 5
US-08-913-880C-13

; Sequence 13, Application US/08913880C
; Patent No. 6372225

; GENERAL INFORMATION:

; APPLICANT: MATSUDA, Morihito

; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS

; FILE REFERENCE: 216-380P

; CURRENT APPLICATION NUMBER: US/08/913, 880C

; CURRENT FILING DATE: 1997-11-24

; NUMBER OF SEQ ID NOS: 17

; SEQ ID NO 13

; LENGTH: 865

; TYPE: PRY

; ORGANISM: Clostridium tetani

; FEATURE:

; OTHER INFORMATION: Amino Acids 451 to 1315 of SEQ ID NO: 1

US-08-913-880C-13

Query Match 99.8%; Score 2451; DB 4; Length 865;
Best Local Similarity 100.0%; Pred. No. 2.7e-189;

Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2 VFSPIPFYSKNDLCWVNEEDIDVILKSTILNDINNDIISDGSFNSVITYPPAQ 61
Db 404 VFSPIPFYSKNDLCWVNEEDIDVILKSTILNDINNDIISDGSFNSVITYPPAQ 463
QY 62 LVPGINGKAHLVNNESSEVIYHKAMDEIYNDMPNFVSWLRPKVSASHLEOYGTNE 121
Db 464 LVPGINGKAHLVNNESSEVIYHKAMDEIYNDMPNFVSWLRPKVSASHLEOYGTNE 523
QY 122 YSISSMKKHSLSIGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPDKFNAYLANKWVF 181
Db 524 YSISSMKKHSLSIGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPDKFNAYLANKWVF 583
QY 182 ITTNDRLSSANLYINGVLGMSAETTGAIREDNNITLKDRCNNNNQYVSIDKFRIFC 241
Db 584 ITTNDRLSSANLYINGVLGMSAETTGAIREDNNITLKDRCNNNNQYVSIDKFRIFC 643
QY 242 KALNPKEIEKLYTSITFLRDFWGNPLRYDTEYLLPVASSSKDVOLKNTIDVWYLTN 301

```

```

Db 644 KALNPKEIEKLYTSITFLRDFWGNPLRYDTEYLLPVASSSKDVOLKNTIDVWYLTN 703
QY 302 APSYTNGLKLNYYRRLYNGLKFTIKRYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 361
Db 704 APSYTNGLKLNYYRRLYNGLKFTIKRYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 763
QY 362 DGNAFNNLDRILRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 421
Db 764 DGNAFNNLDRILRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 823
QY 422 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 463
Db 824 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 865

```

RESULT 6
US-08-913-880C-12

; Sequence 12, Application US/08913880C
; Patent No. 6372225

; GENERAL INFORMATION:

; APPLICANT: MATSUDA, Morihito

; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS

; FILE REFERENCE: 216-380P

; CURRENT APPLICATION NUMBER: US/08/913, 880C

; CURRENT FILING DATE: 1997-11-24

; NUMBER OF SEQ ID NOS: 17

; SEQ ID NO 12

; LENGTH: 866

; TYPE: PRY

; ORGANISM: Clostridium tetani

; FEATURE:

; OTHER INFORMATION: Amino Acids 450 to 1315 of SEQ ID NO: 1

US-08-913-880C-12

Query Match 99.8%; Score 2451; DB 4; Length 866;
Best Local Similarity 100.0%; Pred. No. 2.7e-189;

Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2 VFSPIPFYSKNDLCWVNEEDIDVILKSTILNDINNDIISDGSFNSVITYPPAQ 61
Db 405 VFSPIPFYSKNDLCWVNEEDIDVILKSTILNDINNDIISDGSFNSVITYPPAQ 464
QY 62 LVPGINGKAHLVNNESSEVIYHKAMDEIYNDMPNFVSWLRPKVSASHLEOYGTNE 121
Db 465 LVPGINGKAHLVNNESSEVIYHKAMDEIYNDMPNFVSWLRPKVSASHLEOYGTNE 524
QY 122 YSISSMKKHSLSIGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPDKFNAYLANKWVF 181
Db 525 YSISSMKKHSLSIGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPDKFNAYLANKWVF 584
QY 182 ITTNDRLSSANLYINGVLGMSAETTGAIREDNNITLKDRCNNNNQYVSIDKFRIFC 241
Db 585 ITTNDRLSSANLYINGVLGMSAETTGAIREDNNITLKDRCNNNNQYVSIDKFRIFC 644
QY 242 KALNPKEIEKLYTSITFLRDFWGNPLRYDTEYLLPVASSSKDVOLKNTIDVWYLTN 301
Db 645 KALNPKEIEKLYTSITFLRDFWGNPLRYDTEYLLPVASSSKDVOLKNTIDVWYLTN 704
QY 302 APSYTNGLKLNYYRRLYNGLKFTIKRYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 361
Db 705 APSYTNGLKLNYYRRLYNGLKFTIKRYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 764
QY 362 DGNAFNNLDRILRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 421
Db 765 DGNAFNNLDRILRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 824
QY 422 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 463
Db 825 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 866

```

```

RESULT 7
US-08-913-880C-11
; Sequence 11, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihito
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 11
; LENGTH: 874
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 442 to 1315 of SEQ ID NO: 1
US-08-913-880C-11

```

```

Query Match          99.8%; Score 2451; DB 4; Length 874;
Best Local Similarity 100.0%; Pred. No. 2.8e-189;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 VSTPIPFYSKRLDCWVNEEDIDVILKSTLINDINDIISDISGFSNVITTPDAQ 61
DB 413 VSTPIPFYSKRLDCWVNEEDIDVILKSTLINDINDIISDISGFSNVITTPDAQ 472
QY 62 LVPINGKAIHLVNNSESEVIVHKAMDIEYNDMFNFTVSFWLRPKVSASHLEOYGTNE 121
DB 473 LVPINGKAIHLVNNSESEVIVHKAMDIEYNDMFNFTVSFWLRPKVSASHLEOYGTNE 532
QY 122 YSIISMKKHSLSIGSGWSVSLKGNLITWTLKDSAGEVROITFRDLPDKFNAYLANKWVF 181
DB 533 YSIISMKKHSLSIGSGWSVSLKGNLITWTLKDSAGEVROITFRDLPDKFNAYLANKWVF 592
QY 182 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFC 241
DB 593 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFC 652
QY 242 KALNPKIEKLYTSYSTFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNITDYMYLTN 301
DB 653 KALNPKIEKLYTSYSTFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNITDYMYLTN 712
QY 302 APSYNGKLNIIYRRLYNGKFLIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 361
DB 713 APSYNGKLNIIYRRLYNGKFLIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 772
QY 362 DGNAPNNLDRILRVGNVAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLSGLVGTNHG 421
DB 773 DGNAPNNLDRILRVGNVAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLSGLVGTNHG 832
QY 422 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWIND 463
DB 833 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWIND 874

```

```

RESULT 8
US-08-913-880C-10
; Sequence 10, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihito
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 10
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Clostridium tetani

```

```

; FEATURE:
; OTHER INFORMATION: Amino Acids 441 to 1315 of SEQ ID NO: 1
US-08-913-880C-10

```

```

Query Match          99.8%; Score 2451; DB 4; Length 875;
Best Local Similarity 100.0%; Pred. No. 2.8e-189;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 VSTPIPFYSKRLDCWVNEEDIDVILKSTLINDINDIISDISGFSNVITTPDAQ 61
DB 414 VSTPIPFYSKRLDCWVNEEDIDVILKSTLINDINDIISDISGFSNVITTPDAQ 473
QY 62 LVPINGKAIHLVNNSESEVIVHKAMDIEYNDMFNFTVSFWLRPKVSASHLEOYGTNE 121
DB 474 LVPINGKAIHLVNNSESEVIVHKAMDIEYNDMFNFTVSFWLRPKVSASHLEOYGTNE 533
QY 122 YSIISMKKHSLSIGSGWSVSLKGNLITWTLKDSAGEVROITFRDLPDKFNAYLANKWVF 181
DB 534 YSIISMKKHSLSIGSGWSVSLKGNLITWTLKDSAGEVROITFRDLPDKFNAYLANKWVF 593
QY 182 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFC 241
DB 594 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFC 653
QY 242 KALNPKIEKLYTSYSTFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNITDYMYLTN 301
DB 654 KALNPKIEKLYTSYSTFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNITDYMYLTN 713
QY 302 APSYNGKLNIIYRRLYNGKFLIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 361
DB 714 APSYNGKLNIIYRRLYNGKFLIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 773
QY 362 DGNAPNNLDRILRVGNVAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLSGLVGTNHG 421
DB 774 DGNAPNNLDRILRVGNVAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLSGLVGTNHG 833
QY 422 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWIND 463
DB 834 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWIND 875

```

```

RESULT 9
US-08-913-880C-1
; Sequence 1, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihito
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 1
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-08-913-880C-1

```

```

Query Match          99.8%; Score 2451; DB 4; Length 1315;
Best Local Similarity 100.0%; Pred. No. 5e-189;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 VSTPIPFYSKRLDCWVNEEDIDVILKSTLINDINDIISDISGFSNVITTPDAQ 61
DB 854 VSTPIPFYSKRLDCWVNEEDIDVILKSTLINDINDIISDISGFSNVITTPDAQ 913
QY 62 LVPINGKAIHLVNNSESEVIVHKAMDIEYNDMFNFTVSFWLRPKVSASHLEOYGTNE 121
DB 914 LVPINGKAIHLVNNSESEVIVHKAMDIEYNDMFNFTVSFWLRPKVSASHLEOYGTNE 973
QY 122 YSIISMKKHSLSIGSGWSVSLKGNLITWTLKDSAGEVROITFRDLPDKFNAYLANKWVF 181

```

Db 974 YSIISMKKHSLSIGGWSVSLKGNLIMTLKDSAGEVROITPRDLDPDKFNAYLANKWVF 1033

QY 182 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKDRCNNNNQVYSIDKFRIFC 241

Db 1034 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKDRCNNNNQVYSIDKFRIFC 1093

QY 242 KALNPKIEIKLYSYSLITFLRDPWGNPLRDEYLLIPVASSSKDVOLKNTIDYMYLTN 301

Db 1094 KALNPKIEIKLYSYSLITFLRDPWGNPLRDEYLLIPVASSSKDVOLKNTIDYMYLTN 1153

QY 302 APSYNGKLNYYRRLYNGLKFTIKRTPNNEIDSFVKSDFIKLYSYNNNEHIVGPK 361

Db 1154 APSYNGKLNYYRRLYNGLKFTIKRTPNNEIDSFVKSDFIKLYSYNNNEHIVGPK 1213

QY 362 DGNAPNNIDRLIRVGNAPGIPLYKKMEAVKRLDKTYSVOLKLYDDKNASLGIVGTHNG 421

Db 1214 DGNAPNNIDRLIRVGNAPGIPLYKKMEAVKRLDKTYSVOLKLYDDKNASLGIVGTHNG 1273

QY 422 QIGNDPNRDILIASNWFNHLKDKILGCDWYFVPTDEGWTND 463

Db 1274 QIGNDPNRDILIASNWFNHLKDKILGCDWYFVPTDEGWTND 1315

RESULT 10

US-08-668-381A-5

; Sequence 5, Application US/08668381A

; Patent No. 5780024

; GENERAL INFORMATION:

; APPLICANT: Brown, Robert H.

; APPLICANT: Fishman, Paul S.

; APPLICANT: Francis, Jonathan W.

; APPLICANT: Hosler, Betsy A.

; TITLE OF INVENTION: SUPEROXIDE DISMUTASE/TETANUS TOXIN

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/668,381A

; FILING DATE: 21-JUN-1996

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/000,473

; FILING DATE: 23-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Paul T.

; REGISTRATION NUMBER: 30,164

; REFERENCE/DOCKET NUMBER: 00786/269001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 618 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-668-381A-5

Query Match 99.3%; Score 2440; DB 1; Length 618;

Best Local Similarity 99.6%; Pred. No. 1.3e-188;

Matches 460; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VSTPPEPSYKNDLDCWVNEEDIDVILKKSTLLNDINDNDISDGSFSSVITYTPDAQ 61

Db 157 VSTPPEPSYKNDLDCWVNEEDIDVILKKSTLLNDINDNDISDGSFSSVITYTPDAQ 216

QY 62 LVPGLNGKAIHLYNNSSSPVIVKAMDIEYNOMFNFTYSFWLRPKVSASHLEQGTNE 121

Db 217 LVPGLNGKAIHLYNNSSSPVIVKAMDIEYNOMFNFTYSFWLRPKVSASHLEQGTNE 276

QY 122 YSIISMKKHSLSIGGWSVSLKGNLIMTLKDSAGEVROITPRDLDPDKFNAYLANKWVF 181

Db 277 YSIISMKKHSLSIGGWSVSLKGNLIMTLKDSAGEVROITPRDLDPDKFNAYLANKWVF 336

QY 182 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKDRCNNNNQVYSIDKFRIFC 241

Db 337 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKDRCNNNNQVYSIDKFRIFC 396

QY 242 KALNPKIEIKLYSYSLITFLRDPWGNPLRDEYLLIPVASSSKDVOLKNTIDYMYLTN 301

Db 397 KALNPKIEIKLYSYSLITFLRDPWGNPLRDEYLLIPVASSSKDVOLKNTIDYMYLTN 456

QY 302 APSYNGKLNYYRRLYNGLKFTIKRTPNNEIDSFVKSDFIKLYSYNNNEHIVGPK 361

Db 457 APSYNGKLNYYRRLYNGLKFTIKRTPNNEIDSFVKSDFIKLYSYNNNEHIVGPK 516

QY 362 DGNAPNNIDRLIRVGNAPGIPLYKKMEAVKRLDKTYSVOLKLYDDKNASLGIVGTHNG 421

Db 517 DGNAPNNIDRLIRVGNAPGIPLYKKMEAVKRLDKTYSVOLKLYDDKNASLGIVGTHNG 576

QY 422 QIGNDPNRDILIASNWFNHLKDKILGCDWYFVPTDEGWTND 463

Db 577 QIGNDPNRDILIASNWFNHLKDKILGCDWYFVPTDEGWTND 618

RESULT 11

US-07-618-312A-4

; Sequence 4, Application US/07618312A

; Patent No. 5389540

; GENERAL INFORMATION:

; APPLICANT: Makoff Dr. Andrew J

; APPLICANT: Romanos Dr. Michael A

; APPLICANT: Clare Dr. Jeffrey J

; APPLICANT: Fairweather Dr. Neil F

; TITLE OF INVENTION: VACCINES

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: 14th Floor

; STREET: 2200 Clarendon Boulevard,

; CITY: Arlington,

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22201

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/618,312A

; FILING DATE: 19910516

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 8926832.0

; FILING DATE: 28-NOV-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9006097.1

; FILING DATE: 17-MAR-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Crawford Mr. Arthur R

; REGISTRATION NUMBER: 25,327

; REFERENCE/DOCKET NUMBER: 510-51

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 0101 703 8750400

TELEFAX: 0101 703 5253468
 TELEEX: 200797 NIXN UR
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 452 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-618-312A-4

Query Match 97.4%; Score 2393; DB 1; Length 452;
 Best Local Similarity 100.0%; Pred. No. 5.1e-185;
 Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 KNDQWVNEEDIDIVILKSTLTLLNDIINDIISDGFNSSVITPPDAOLVPGINGKAH 72
 DB 2 KNDQWVNEEDIDIVILKSTLTLLNDIINDIISDGFNSSVITPPDAOLVPGINGKAH 61
 QY 73 LVNSESSEVIYHKAMDIENDMFNFTVSFWLRVPKVSASHLEQYGTNEYSTISSMKKHS 132
 DB 62 LVNSESSEVIYHKAMDIENDMFNFTVSFWLRVPKVSASHLEQYGTNEYSTISSMKKHS 121
 QY 133 LSTGSGWSVSLKGNLITWLKDSAGEVROITPRDLPDKFNAYLANRWFTTTNDRLSA 192
 DB 122 LSTGSGWSVSLKGNLITWLKDSAGEVROITPRDLPDKFNAYLANRWFTTTNDRLSA 181
 QY 193 NLYINGVLMSAEITGLAIREDDNITTLKDRCNNNNOYVSIDKFRIFCKALNPKREIEKL 252
 DB 182 NLYINGVLMSAEITGLAIREDDNITTLKDRCNNNNOYVSIDKFRIFCKALNPKREIEKL 241
 QY 253 YTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNTDYMYLTNAPSYTNGKLN 312
 DB 242 YTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNTDYMYLTNAPSYTNGKLN 301
 QY 313 YVRRLYNGLKFIIRKRYTPNNEIDSFYKSGDFIKLYSYNNNEHIVGPKDGNAFNNLDRI 372
 DB 302 YVRRLYNGLKFIIRKRYTPNNEIDSFYKSGDFIKLYSYNNNEHIVGPKDGNAFNNLDRI 361
 QY 373 LRVGYNAPGIPLYKKMAVLRDLKTYSVOLKLYDDKNASLGLVGTNGOIGNDPNRDL 432
 DB 362 LRVGYNAPGIPLYKKMAVLRDLKTYSVOLKLYDDKNASLGLVGTNGOIGNDPNRDL 421
 QY 433 IASNWFNHLKDKITLGGDWYFVPTDEGWTND 463
 DB 422 IASNWFNHLKDKITLGGDWYFVPTDEGWTND 452

RESULT 12

US-08-280-228-4
 Sequence 4, Application US/08280228

GENERAL INFORMATION:
 PATENT NO. 5571694
 APPLICANT: Makoff Dr, Andrew J
 APPLICANT: Romanos Dr, Michael A
 APPLICANT: Clare Dr, Jeffrey J
 APPLICANT: Fairweather Dr, Neil F
 TITLE OF INVENTION: VACCINES
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: NIXON & VANDERHAYE P.C.
 STREET: 1100 No. 5571694th Glebe Road
 CITY: Arlington,
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22201-4714
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/280,228
 FILING DATE: 25-JUL-1994

CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/618,312
 FILING DATE: 27-NOV-1990
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 8926832.0
 FILING DATE: 28-NOV-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9006097.1
 FILING DATE: 17-MAR-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Wilson, Mary J.
 REGISTRATION NUMBER: 32,955
 REFERENCE/DOCKET NUMBER: 117-163
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4000
 TELEFAX: (703) 816-4100
 TELEEX: 200797 NIXN UR
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 452 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-280-228-4

Query Match 97.4%; Score 2393; DB 1; Length 452;
 Best Local Similarity 100.0%; Pred. No. 5.1e-185;
 Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 KNDQWVNEEDIDIVILKSTLTLLNDIINDIISDGFNSSVITPPDAOLVPGINGKAH 72
 DB 2 KNDQWVNEEDIDIVILKSTLTLLNDIINDIISDGFNSSVITPPDAOLVPGINGKAH 61
 QY 73 LVNSESSEVIYHKAMDIENDMFNFTVSFWLRVPKVSASHLEQYGTNEYSTISSMKKHS 132
 DB 62 LVNSESSEVIYHKAMDIENDMFNFTVSFWLRVPKVSASHLEQYGTNEYSTISSMKKHS 121
 QY 133 LSTGSGWSVSLKGNLITWLKDSAGEVROITPRDLPDKFNAYLANRWFTTTNDRLSA 192
 DB 122 LSTGSGWSVSLKGNLITWLKDSAGEVROITPRDLPDKFNAYLANRWFTTTNDRLSA 181
 QY 193 NLYINGVLMSAEITGLAIREDDNITTLKDRCNNNNOYVSIDKFRIFCKALNPKREIEKL 252
 DB 182 NLYINGVLMSAEITGLAIREDDNITTLKDRCNNNNOYVSIDKFRIFCKALNPKREIEKL 241
 QY 253 YTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNTDYMYLTNAPSYTNGKLN 312
 DB 242 YTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNTDYMYLTNAPSYTNGKLN 301
 QY 313 YVRRLYNGLKFIIRKRYTPNNEIDSFYKSGDFIKLYSYNNNEHIVGPKDGNAFNNLDRI 372
 DB 302 YVRRLYNGLKFIIRKRYTPNNEIDSFYKSGDFIKLYSYNNNEHIVGPKDGNAFNNLDRI 361
 QY 373 LRVGYNAPGIPLYKKMAVLRDLKTYSVOLKLYDDKNASLGLVGTNGOIGNDPNRDL 432
 DB 362 LRVGYNAPGIPLYKKMAVLRDLKTYSVOLKLYDDKNASLGLVGTNGOIGNDPNRDL 421
 QY 433 IASNWFNHLKDKITLGGDWYFVPTDEGWTND 463
 DB 422 IASNWFNHLKDKITLGGDWYFVPTDEGWTND 452

RESULT 13

US-07-618-312A-2

Sequence 2, Application US/07618312A
 Patent No. 538540
 GENERAL INFORMATION:
 APPLICANT: Makoff Dr, Andrew J
 APPLICANT: Romanos Dr, Michael A
 APPLICANT: Clare Dr, Jeffrey J
 APPLICANT: Fairweather Dr, Neil F

1 TITLE OF INVENTION: VACCINES
2 NUMBER OF SEQUENCES: 13
3 CORRESPONDENCE ADDRESSES:
4 ADDRESSEE: 14th Floor
5 STREET: 2200 Clarendon Boulevard,
6 CITY: Arlington,
7 STATE: Virginia
8 COUNTRY: U.S.A.
9 ZIP: 22201
10
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Floppy disk
13 OPERATING SYSTEM: IBM PC compatible
14 SOFTWARE: Patentin Release #1.0, Version #1.25
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/07/618,312A
17 FILING DATE: 19910516
18 CLASSIFICATION: 424
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: GB 8926832.0
21 FILING DATE: 28-NOV-1989
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: GB 9006097.1
24 FILING DATE: 17-MAR-1990
25 ATTORNEY/AGENT INFORMATION:
26 NAME: Crawford Mr, Arthur R
27 REGISTRATION NUMBER: 25,327
28 REFERENCE/DOCKET NUMBER: 510-51
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: 0101 703 8750400
31 TELEFAX: 0101 703 5253468
32 TELEX: 200797 NIXN UR
33 INFORMATION FOR SEQ ID NO: 2:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 452 amino acids
36 TYPE: AMINO ACID
37 TOPOLOGY: linear
38 MOLECULE TYPE: protein
39 US-07-618-312A-2

Query Match 96.9%; Score 2380; DB 1; Length 452;
Best Local Similarity 99.6%; Pred. No. 5.7e-184;
Matches 449; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

13 KNLDCWVNEEDIDVILKSTILNDINNDIISDGFNSVITYPDAQLVPGINGKAH 72
2 KNLDCWVNEEDIDVILKSTILNDINNDIISDGFNSVITYPDAQLVPGINGKAH 61
73 LVNNESEVIVHAKMDIEYNDMFNFTVSEWLRVPKVSASHLEQYGTNEYSITSSMKKHS 132
62 LVNNESEVIVHAKMDIEYNDMFNFTVSEWLRVPKVSASHLEQYGTNEYSITSSMKKHS 121
133 LSIISGWSVSLKGNLILWTLKDSAGEVROITFRDLPRKFNAYLANKWFVITITNDRLSA 192
122 LSIISGWSVSLKGNLILWTLKDSAGEVROITFRDLPRKFNAYLANKWFVITITNDRLSA 181
193 NLYINGVLMGSAITIGALIREDDNNTILKIDRCNNNNQYVSIIDKFRIFCKALNKEIEKL 252
182 NLYINGVLMGSAITIGALIREDDNNTILKIDRCNNNNQYVSIIDKFRIFCKALNKEIEKL 241
253 YTSYLSITFIRBDMGNPLRDTREYLLIPVASSSKDVOLKNITDYMVLNAPSITNGKLN 312
242 YTSYLSITFIRBDMGNPLRDTREYLLIPVASSSKDVOLKNITDYMVLNAPSITNGKLN 301
313 YTRRLYNGLKFIKRYTPNNEIDSEVKSDFIKLYSYNNNEHIVGPKGNMFPNNLDRI 372
302 YTRRLYNGLKFIKRYTPNNEIDSEVKSDFIKLYSYNNNEHIVGPKGNMFPNNLDRI 361
373 LVGVNAPRIPILYKKEAVKLDLKYTVOLKLYDKNASLGLVGTNGOIGNDPNRDL 432
362 LVGVNAPRIPILYKKEAVKLDLKYTVOLKLYDKNASLGLVGTNGOIGNDPNRDL 421
433 IASNMVFNLKDKILGCDWYFVPTDEGWTND 463

Db 422 IASNMVFNLKDKILGCDWYFVPTDEGWTND 452

RESULT 14
US-08-280-228-2
Sequence 2, Application US/08280228
Patent No. 5571694
GENERAL INFORMATION:
APPLICANT: Makoff Dr, Andrew J
APPLICANT: Romanos Dr, Michael A
APPLICANT: Clare Dr, Jeffrey J
APPLICANT: Fairweather Dr, Neil F
TITLE OF INVENTION: VACCINES
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 No. 5571694th Glebe Road
CITY: Arlington,
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/280,228
FILING DATE: 25-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/618,312
FILING DATE: 27-NOV-1990
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8926832.0
FILING DATE: 28-NOV-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9006097.1
FILING DATE: 17-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mary J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 117-163
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
TELEX: 200797 NIXN UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 452 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-280-228-2

Query Match 96.9%; Score 2380; DB 1; Length 452;
Best Local Similarity 99.6%; Pred. No. 5.7e-184;
Matches 449; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

13 KNLDCWVNEEDIDVILKSTILNDINNDIISDGFNSVITYPDAQLVPGINGKAH 72
2 KNLDCWVNEEDIDVILKSTILNDINNDIISDGFNSVITYPDAQLVPGINGKAH 61
73 LVNNESEVIVHAKMDIEYNDMFNFTVSEWLRVPKVSASHLEQYGTNEYSITSSMKKHS 132
62 LVNNESEVIVHAKMDIEYNDMFNFTVSEWLRVPKVSASHLEQYGTNEYSITSSMKKHS 121
133 LSIISGWSVSLKGNLILWTLKDSAGEVROITFRDLPRKFNAYLANKWFVITITNDRLSA 192
122 LSIISGWSVSLKGNLILWTLKDSAGEVROITFRDLPRKFNAYLANKWFVITITNDRLSA 181

QY 193 NLXYINGVLMGSAEITGLGAIREDNNITLKDRCNNNOYVSDKFRIFCKALNPKIEKL 252
 DB 182 NLXYINGVLMGSAEITGLGAIREDNNITLKDRCNNNOYVSDKFRIFCKALNPKIEKL 241
 QY 253 YTSYSTIFLDRDFWGNPLRYDTEYLIIPVASSSKDVOLKNTIDYMLTNAPSYTNGKLN 312
 DB 242 YTSYSTIFLDRDFWGNPLRYDTEYLIIPVASSSKDVOLKNTIDYMLTNAPSYTNGKLN 301
 QY 313 YRRLYNGLKFIIKRTYPNNEIDSFVKSDFIKLYSYNNNEHIVGYPKDGNAFNNDRI 372
 DB 302 YRRLYNGLKFIIKRTYPNNEIDSFVKSDFIKLYSYNNNEHIVGYPKDGNAFNNDRI 361
 QY 373 LRVGNAPGIPLYYKKMEAVKLRDLKTYVOLKLYDDKNASLGVTGHNQIGNDPRDIL 432
 DB 362 LRVGNAPGIPLYYKKMEAVKLRDLKTYVOLKLYDDKNASLGVTGHNQIGNDPRDIL 421
 QY 433 IASNNYFNHLKDKILGCDWYFVPTDEGWIND 463
 DB 422 IASNNYFNHLKDKILGCDWYFVPTDEGWIND 452

RESULT 15

US-08-110-786A-8
 ; Sequence 8, Application us/08110786A

; Patent No. 5443966
 ; GENERAL INFORMATION:

APPLICANT: FAIRWEATHER, Neil Fraser

APPLICANT: MAKOFF, Andrew Joseph

TITLE OF INVENTION: Expression of tetanus toxin fragment C

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon & Vanderhye P.C.

STREET: 1100 No. 5443966th Glebe Road

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/110,786A

FILING DATE: 23-AUG-1993 1991

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/777,337

FILING DATE: 29-NOV-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB90/00943

FILING DATE: 20-JUN-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 8914122.0

FILING DATE: 20 June 1989

ATTORNEY/AGENT INFORMATION:

NAME: Mary J. Wilson

REGISTRATION NUMBER: 32,955

REFERENCE/DOCKET NUMBER: 117-134

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100

TELEX: 200797 NIXN UR

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 452 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-110-786A-8

Query Match 95.5%; Score 2345; DB 1; Length 452;

Best Local Similarity 98.4%; Pred. No. 3.8e-181;

Matches 444; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 13 KNLDCWVNEEDIDIVYILKSTILNDINNDIISDGSNSVITYPPDAQVPGINGKAH 72
 DB 2 KNLDCWVNEEDIDIVYILKSTILNDINNDIISDGSNSVITYPPDAQVPGINGKAH 61
 QY 73 LVNNESSFVIVHKAMDIENDFNPFVSWLRVPKVSASHLEQYGTNEYSSINSSMKHS 132
 DB 62 LVNNESSFVIVHKAMDIENDFNPFVSWLRVPKVSASHLEQYGTNEYSSINSSMKHS 121
 QY 133 LSTGSGWSYSLKGNLWTLKDSAGEVROIFRDLDPKFNALANKWPTITNDRLSSA 192
 DB 122 LSTGSGWSYSLKGNLWTLKDSAGEVROIFRDLDPKFNALANKWPTITNDRLSSA 181
 QY 193 NLXYINGVLMGSAEITGLGAIREDNNITLKDRCNNNOYVSDKFRIFCKALNPKIEKL 252
 DB 182 NLXYINGVLMGSAEITGLGAIREDNNITLKDRCNNNOYVSDKFRIFCKALNPKIEKL 241
 QY 253 YTSYSTIFLDRDFWGNPLRYDTEYLIIPVASSSKDVOLKNTIDYMLTNAPSYTNGKLN 312
 DB 242 YTSYSTIFLDRDFWGNPLRYDTEYLIIPVASSSKDVOLKNTIDYMLTNAPSYTNGKLN 301
 QY 313 YRRLYNGLKFIIKRTYPNNEIDSFVKSDFIKLYSYNNNEHIVGYPKDGNAFNNDRI 372
 DB 302 YRRLYNGLKFIIKRTYPNNEIDSFVKSDFIKLYSYNNNEHIVGYPKDGNAFNNDRI 361
 QY 373 LRVGNAPGIPLYYKKMEAVKLRDLKTYVOLKLYDDKNASLGVTGHNQIGNDPRDIL 432
 DB 362 LRVGNAPGIPLYYKKMEAVKLRDLKTYVOLKLYDDKNASLGVTGHNQIGNDPRDIL 421
 QY 433 IASNNYFNHLKDKILGCDWYFVPTDEGWIND 463
 DB 422 IASNNYFNHLKDKILGCDWYFVPTDEGWIND 452

Search completed: December 1, 2002, 11:25:31
 Job time : 29 secs


```

QY 229 NOVVIDKFRICAKALNEKEIEKLYSTFLKDFGNPLRDEYLLIPVASSKDV 288
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1063 HRYIMIKFENLEKELNEKEIKLDYDQNSGILKDFWGDYLDKPYMLNLDYKPY 1122
QY 289 QLK--ITDYATLNAPSYTNGKLNYYR-RLYNGKFLIRYPPNNEIDSPVSGDFIK 345
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1123 DVNNVIGGYWL-KGPRGSVWTTNLYNSSLYRGTFIKRYKVASGNK-DNIVRNDRY 1180
QY 346 LYVSNNNEHLYGVKDNNAENLDRILRVGNAPGIPLYKMEAVKLRDK----- 397
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1181 INVYKKEY-----RIATNASQAGEKLSLEIPDGNLSQVVM 1222
QY 398 -----TYSVOLKLYDDKNASGLGVTHNGQICNDPNRDLIASNNYFNHLK--DKIL 447
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1223 KSKNDGTTNCKMNLQNNNGNDIGIFGH-QFNNIK--LVASNNYNRQIERSSRTL 1277
QY 448 GCDWYFVPTDEGM 460
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1278 GCSWEFIPVDDGM 1290

```

RESULT 2

```

US-09-288-326-9
/ Sequence 9, Application US/09288326
/ Patent No. US20010018049A1
/ GENERAL INFORMATION:
/ APPLICANT: Kel Roger Aoki
/ APPLICANT: George Sachs
/ TITLE OF INVENTION: Method and Compositions for the
/ TITLE OF INVENTION: Treatment of Pancreatitis
/ FILE REFERENCE: 17282
/ CURRENT APPLICATION NUMBER: US/09/288,326
/ CURRENT FILING DATE: 1999-04-08
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 9
/ LENGTH: 425
/ TYPE: PRY
/ ORGANISM: Clostridium botulinum
/ US-09-288-326-9

```

```

Query Match 25.5%; Score 627.5; DB 10; Length 425;
Best Local Similarity 33.3%; Pred. No. 4,4e-40;
Matches 153; Conservative 84; Mismatches 149; Indels 73; Gaps 16;

```

```

QY 32 STLINDINNDIISDISGNSVITYPDAOLVPGINGKAIHLVNNESSEVIYKAMDIEY 91
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 TSLINLRYESNHLIDLSRYASKINIGSKVNFDP-IDKNQIQLFNLESSKEVLKNAIVY 63
QY 92 NDMFNNTFVEMLRVPK-----VSASHLEQYGTNEYSTIISMKKHSLSISGWSYSLKGN 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 NSMTEENSTFWMIRIPYFNSISLN-----NEYTIINCENN-----SGMKVSLNAYGE 111
QY 148 LIWTLKDSAGEVROITER-----DLDPKFNAVYLANWKVFTITNDRLSSANLYINGVLMG 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 112 IIMTLQDOTQKQVYKYSOMINISD-----YI-NMWIFVITNNRLGNSRYINGNLID 166
QY 203 SAETTGAGAIREDNNITLKDRCNNNNQYVSIKFRIFECALNEKEIKLYTSSTLFL 262
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 167 OKPISLNGNTHASNNIMFKLDGCDTHRYIMIKYFLDEKELNEKEIKLDYDQNSGIL 226
QY 263 RDEFGNPLRYDEYLLIPVASSKDVQKLN--ITDYATLNAPSYTNGKLNYYR-RLYN 319
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 227 KDEFGDLYQYKPYMLNLDYKPYNNVYDNNVNGIRGYWL-KGPRGSVWTTNLYNSSLYR 285
QY 320 GLKFIIRKYPNNEIDSPVSGDFIKLYSTNNNEHLYGVKDGNAFPNNDRIILRVGNA 379
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 286 GTFKFIIRKVASGNK-DNIVRNDRYINVYKKEY-----RIATNA 326
QY 380 PGIFLYKMEAVKLRDK-----TYSVOLKLYDDKNASGLGVTHNGQI 423
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 327 SQAGCEKLSLEIPDGNLSQVVMKSKNDGITNCKMNLQNNNGNDIGIFGH--QF 384

```

```

QY 424 GNDPNRDLIASNNYFNHLK--DKILGCDWYFVPTDEGM 460
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 385 NNIAK--LVASNNYNRQIERSSRTLCSWEFIPVDDGM 420

```

RESULT 3

```

US-08-981-087A-1
/ Sequence 1, Application US/08981087A
/ Patent No. US20020081304A1
/ GENERAL INFORMATION:
/ APPLICANT: Elmore, Michael J.
/ APPLICANT: Mauchline, Margaret L.
/ APPLICANT: Minton, Nigel P.
/ APPLICANT: Pasechnik, Vladimir A.
/ APPLICANT: Tibball, Richard W.
/ TITLE OF INVENTION: TYPE F BOTULINUM TOXIN AND USE THEREOF
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: NIXON & VANDERHAYE P.C.
/ STREET: 1100 No. US20020081304A1th Glebe Rd. 8th floor
/ CITY: Arlington
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22201-4741
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/981,087A
/ FILING DATE: 27-MAY-1998
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/GB96/01409
/ FILING DATE: 12-JUN-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9511909.5
/ FILING DATE: 12-JUN-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Crawford, Arthur R.
/ REGISTRATION NUMBER: 25,327
/ REFERENCE/DOCKET NUMBER: 124-688
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 703-816-4100
/ TELEFAX: 703-816-4100
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 431 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-08-981-087A-1

```

```

Query Match 22.5%; Score 553; DB 8; Length 431;
Best Local Similarity 30.0%; Pred. No. 1.8e-34;
Matches 138; Conservative 92; Mismatches 152; Indels 78; Gaps 18;

```

```

QY 29 LKSTIINDINNDIISDISGNSVITYPDAOLVPGINGKAIHLVNNESSEVIYKAMDIEY 88
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 19 IKDSIINDMRENNKFLDISGNSISNGDVYIY-STNRQOFGIYSSKPSVINAQND 77
QY 89 IEYNDMENFTVSWLRLVPKVASASHLEQY-TNEYSTIISMKKHSLSISGWSYSLKGN 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 78 IYNGRQNFSTISFWMIRPK-----YFNKVNLNNEYTIIDCIARNN-----SGMKISLANK 129
QY 148 LIWTLKDSAGEVROITER-----DLDPKFNAVYLANWKVFTITNDRLSSANLYINGVLMG 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 IIMTLQDOTAGNNQKLVYTYOMISID-----YI-NKWIFVITNNRLGNSRYINGNLID 184
QY 203 SAETTGAGAIREDNNITLKDRCNNNNQYVSIKFRIFECALNEKEIKLYTSSTLFL 262
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```


Db 130 IITWIDDTAGNNOKL 144

RESULT 6

US-08-981-087A-4
; Sequence 4, Application US/08981087A
; Patent No. US20020081304A1

GENERAL INFORMATION:

APPLICANT: Elmore, Michael J.

APPLICANT: Mauchline, Margaret L.

APPLICANT: Minton, Nigel P.

APPLICANT: Paschuk, Vladimir A.

APPLICANT: Tibbalt, Richard W.

TITLE OF INVENTION: TYPE F BOTULINUM TOXIN AND USE THEREOF

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESSEE: NIXON & VANDERHAYE P.C.

STREET: 1100 No. US20020081304A1th Glabe Rd. 8th floor

CITY: Arlington

STATE: VA

COUNTRY: USA

ZIP: 22201-4741

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/981,087A

FILING DATE: 27-MAY-1998

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB96/01409

FILING DATE: 12-JUN-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9511909.5

FILING DATE: 12-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Crawford, Arthur R.

REGISTRATION NUMBER: 25,327

REFERENCE/DOCKET NUMBER: 124-688

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-816-4100

TELEFAX: 703-816-4100

INFORMATION FOR SEQ. ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 143 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-981-087A-4

Query Match 5.6%; Score 137; DB 8; Length 143;
Best Local Similarity 23.4%; Pred. No. 0.00089;
Matches 40; Conservative 32; Mismatches 47; Indels 52; Gaps 8;

QY 311 NIYYR-RLVNGLFLIR--YTPNNEIDSFYKSGD-----IKLV--VSNNNE 354
Db 1 NITSNRLTYGVEYITIKKNSDTISNTDFYRKNDLAYINVDVDEYRLADISIAKE 60
QY 355 HAVGYPKGNANFNLDRL--RVGYNAPGIPLYKKMEAVKRLDKTYSVOLKLYDDKNA 411
Db 61 KIILKLTSSNNSLGLIWDISGNCC-----TNNFNNNGG 98
QY 412 SLGLVGHNGOIGNDPRNDILASNWFNHLKDKIL--GCDWYFVPPDEGW 460
Db 99 NIGLGHFSNN-----LVASSWYNNIRKNTSSNGCFWSPISKHGW 140

RESULT 7
US-09-797-097-2
; Sequence 2, Application US/09797097

; Patent No. US20020058312A1
; GENERAL INFORMATION:
; APPLICANT: All, Benjamin
; APPLICANT: Howard, John
; TITLE OF INVENTION: Expression Cassettes and Methods of Delivery of Animal Vaccine
; FILE REFERENCE: P002460U1
; CURRENT APPLICATION NUMBER: US/09/797,097
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 08/529,006
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 1447
; TYPE: PRT
; ORGANISM: Transmissible Gastroenteritis Virus Spike protein
US-09-797-097-2

Query Match 5.0%; Score 124; DB 10; Length 1447;
Best Local Similarity 16.9%; Pred. No. 0.18;
Matches 89; Conservative 76; Mismatches 162; Indels 200; Gaps 21;

QY 30 KKSFLINDINDIDISIGFSSVITPDQVLPGINGKAIHLVNNESSEVIVHAMD 89
Db 354 KGAIVFSLNTTGGVLEISCYNDIV-----SPSSFSYGEMPS 391
QY 90 EYND-----MENNFTVSFWLRV---KVSASHLEOYGTNEYSISSMKKHSLSIGS 138
Db 392 GVTDSRRCYVLYNCTALKYLSLTPPYKEIAMSCKGHYINGYNFSPIDCIS---- 447
QY 139 WSVSLKGNLWTLKDSAGEVROITFRDLPKFNAYLAKWVFITTN---DLSSANLY 195
Db 448 FNLFTGDSDFEWTI-----ATSYTEALVOVENAATKVTYCNSTY 487
QY 196 INGVL-----MGSAEI-----TGCAIRED----- 215
Db 488 VNNIKCSQTLANLNGFYPVSSVEGLVYKSVYLLPFTHTITVNTTIGLGMRSVGP 547
QY 216 ---NNTLKDRCNNNNQVYSIDKFRICKALNKEIEKLYSYLSITFLDEMGNPL 270
Db 548 IASTLSNLTLPQDNNNTDVCIRSDQFSV-----YVHSTCTSLMDNVP 591
QY 271 RYDTEYLYLVASSSKQVQDKNITDYM-----YLTNAPSTNGKLNYYRLYN----- 319
Db 592 K-----RNCITDVLATPAVTKTCTCPSEPKLNNYL--TNNKFLSL 630
QY 320 ---GLKFLIKRYPNNEIDSFYKSGDFIKLYVSYNNMEHIVYKPDGNAFNLDRLR 374
Db 631 SPVGNACKRFDVAKRTFN--DQVRS-----LVYIEEGDNIVGVPDNGSLDLS- VLH 682
QY 375 VG-----YVNPAGIPLYKKMEAVKRLDKTYSVOLKLYDDKNAISGLV----- 416
Db 683 IDSCIDYNYGKRGVGIIRQTNRTLLSGLYYTSGLDGLFKKWSVDVIVSYVPCDVSAQ 742
QY 417 -----GTHNGOIGNDPRNDILASNWFNHLKDKILG 448
Db 743 AAVIDGTVGAI-TSINSELLGLTHWTTTPNPFYYSIVTYNTDRTRG 788

RESULT 8
US-09-801-574-57
; Sequence 57, Application US/09801574
; Patent No. US20020081592A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Peljing Jeremy
; APPLICANT: Page, David C.
; TITLE OF INVENTION: Reproduction-Specific Genes
; FILE REFERENCE: 0399.2007-002
; CURRENT APPLICATION NUMBER: US/09/801,574
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/187,518
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: 60/261,557

; PRIOR FILING DATE: 2001-01-12
 ; NUMBER OF SEQ ID NOS: 90
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 57
 ; LENGTH: 2789
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-801-574-57

Query Match 4.7%; Score 116.5; DB 10; Length 2789;
 Best Local Similarity 22.2%; Pred. No. 1.6;
 Matches 111; Conservative 79; Mismatches 164; Indels 147; Gaps 30;

QY 3 FSPPIPSYKNDLDCWDNEEDIDVILKSTILMLDINNDIISDISGNSVYTPDAQ- 61
 DB 1806 FSLSVPTGCVNEG---DSLEDE-ILRKSTLK-----LINVCGDSPVHSYPGKD 1853
 QY 62 ---LVGINKALHLVNNSSSEY-----IVHKAMDIEYMDNNFVSR----- 102
 DB 1854 HLMIIEMISSKVFNNNEAVRKISLYGLEHLEFPAKMLVWKEKQSFSSKYSQKRD 1913
 QY 103 ---WLRVRYKVSASHLEQ-----YGTNEYSISSMK-KHSLISGWSVS 142
 DB 1914 EERLLRYNKCASFSLKITYDTLSKDLNNEPIISGLEEDTIIASRKSHPINEA---TIS 1970
 QY 143 LK---GNL-----ITLKDSA-----GEYRQITFR-----DLPDF----- 171
 DB 1971 IENSKRNSNLAHPDICISILDOAEFADLKLQDLTLRCTDHLFLKRYFQMLQDNM 2030
 QY 172 -NAVLANRWFFITND-----RLSSANLYINGVMSGSEITGLAIRDNNITLKL 223
 DB 2031 DNFIIEENVLDVIVINSHAIILKPEALIMYLE-IVWSEITLHFL-----KNSIAR 2084
 QY 224 RCNNNNQVYSIDKFR---IFCKALNPEIEKLYSYLSITFLRD-----FWGNPLRYDT 274
 DB 2085 K-----QFRGMLWEDLSLP-ELVQCEKMASSEFLKNSNDVCIM----- 2125
 QY 275 EYLLIPVASSKNOVKNTIDYMYLTINAPSYTKLNTIYRRL--YNGLFILKRYPPN 332
 DB 2126 --KVITTAVSELRKDDITICKINEAVNC-SYA--IHLSRELOELSEIKKLK--SKY 2177
 QY 333 EIDSPVKSQDFTKLYVSYNNNEHIVGYPKDNAPNNLRLIRGYNAPGPIYKMEAVK 392
 DB 2178 FISTYI--DFVPYIASINVGSTVELEYNYNOFSTL--LKNVMSAPKRDGLKMAHIRK 2231
 QY 393 LMDLTIYSQVLKLYDKNASL 413
 DB 2232 V--MKTIE-HMKMICYKNAEL 2249

RESULT 9

US-09-848-834A-14
 ; Sequence 14, Application US/09848834A
 ; Patent No. US20020076416A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aphron Corporation
 ; TITLE OF INVENTION: Chimeric Peptide Immunogens
 ; FILE REFERENCE: 1102865-0047
 ; CURRENT APPLICATION NUMBER: US/09/848,834A
 ; CURRENT FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: 60/202,328
 ; PRIOR FILING DATE: 2000-05-05
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 14
 ; LENGTH: 37
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the G
 ; OTHER INFORMATION: nrh hormone linked by a spacer to amino acid sequence 947-967 of
 ; OTHER INFORMATION: the Tetanus toxoid precursor (Tentoxylisin)
 ; NAME/KEY: MOD_RES

; LOCATION: (1)..(1)
 ; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (1)..(10)
 ; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (11)..(16)
 ; OTHER INFORMATION: Spacer peptide
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (17)..(37)
 ; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor
 ; OTHER INFORMATION: (Tentoxylisin)
 US-09-848-834A-14

Query Match 4.6%; Score 114; DB 10; Length 37;
 Best Local Similarity 95.5%; Pred. No. 0.0081;
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 94 MFNNFTVSFWLRVRYKVSASHLE 115
 DB 16 LFNFTVSFWLRVRYKVSASHLE 37

RESULT 10

US-09-848-834A-18
 ; Sequence 18, Application US/09848834A
 ; Patent No. US20020076416A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aphron Corporation
 ; TITLE OF INVENTION: Chimeric Peptide Immunogens
 ; FILE REFERENCE: 1102865-0047
 ; CURRENT APPLICATION NUMBER: US/09/848,834A
 ; CURRENT FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: 60/202,328
 ; PRIOR FILING DATE: 2000-05-05
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 18
 ; LENGTH: 50
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of hum
 ; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the
 ; OTHER INFORMATION: anus toxoid precursor (Tentoxylisin) protein linked by a spac
 ; OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH
 ; NAME/KEY: MOD_RES
 ; LOCATION: (1)..(11)
 ; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
 ; NAME/KEY: MOD_RES
 ; LOCATION: (50)..(50)
 ; OTHER INFORMATION: Amidated glycine or glycylamide
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (1)..(10)
 ; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (11)..(16)
 ; OTHER INFORMATION: Spacer peptide
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (17)..(37)
 ; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (T
 ; OTHER INFORMATION: oxylysin
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (38)..(41)
 ; OTHER INFORMATION: Spacer peptide
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (42)..(50)
 ; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
 US-09-848-834A-18

Query Match 4.6%; Score 114; DB 10; Length 50;
 Best Local Similarity 95.5%; Pred. No. 0.012;
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 94 MNFTVSWLRVPKVSASHLE 115
 :|||||
 Db 16 LFNFTVSWLRVPKVSASHLE 37

RESULT 11
 US-09-943-548-3
 ; Sequence 3, Application US/09943548
 ; Patent No. US20020042364A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rittershaus, Charles W.
 ; APPLICANT: Thomas, Lawrence J.
 ; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVITY
 ; FILE REFERENCE: TCS-411.1P US-1; TCS-411.1P US-2
 ; CURRENT APPLICATION NUMBER: US/09/943,548
 ; PRIOR FILING DATE: 2001-08-30
 ; PRIOR APPLICATION NUMBER: 08/432,483
 ; PRIOR FILING DATE: 1995-05-01
 ; PRIOR APPLICATION NUMBER: PCT/US96/06147
 ; PRIOR FILING DATE: 1996-05-01
 ; PRIOR APPLICATION NUMBER: 08/945,289
 ; PRIOR FILING DATE: 1997-10-17
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 21
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: helper T cell epitope of tetanus toxin
 US-09-943-548-3

Query Match 4.6%; Score 112; DB 10; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.0054;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 95 FNNFTVSWLRVPKVSASHLE 115
 :|||||
 Db 1 FNNFTVSWLRVPKVSASHLE 21

RESULT 12
 US-09-848-834A-4
 ; Sequence 4, Application US/09848834A
 ; Patent No. US20020076416A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Apton Corporation
 ; TITLE OF INVENTION: Chimeric Peptide Immunogens
 ; FILE REFERENCE: 1102865-0047
 ; CURRENT APPLICATION NUMBER: US/09/848,834A
 ; CURRENT FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: 60/202,328
 ; PRIOR FILING DATE: 2000-05-05
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 21
 ; TYPE: PRT
 ; ORGANISM: Tetanus bacillus
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (1)..(21)
 ; OTHER INFORMATION: Amino acid sequence 947-967 of Tetanus
 ; OTHER INFORMATION: Toxoid Precursor (Tentoxylysin)
 US-09-848-834A-4

Query Match 4.6%; Score 112; DB 10; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.0054;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 95 FNNFTVSWLRVPKVSASHLE 115
 :|||||

Db 1 FNNFTVSWLRVPKVSASHLE 21

RESULT 13
 US-09-848-834A-10
 ; Sequence 10, Application US/09848834A
 ; Patent No. US20020076416A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Apton Corporation
 ; TITLE OF INVENTION: Chimeric Peptide Immunogens
 ; FILE REFERENCE: 1102865-0047
 ; CURRENT APPLICATION NUMBER: US/09/848,834A
 ; CURRENT FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: 60/202,328
 ; PRIOR FILING DATE: 2000-05-05
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 10
 ; LENGTH: 34
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 947-967 of
 ; OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylysin) linked by a spacer to
 ; OTHER INFORMATION: Ino acid sequence 2-10 of the GnRH hormone
 ; NAME/KEY: MOD_RES
 ; LOCATION: (1)..(1)
 ; OTHER INFORMATION: Amidated phenylalanine
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (1)..(21)
 ; OTHER INFORMATION: Amino acids 947-967 of the Tetanus Toxoid Precursor
 ; OTHER INFORMATION: (Tentoxylysin)
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (22)..(25)
 ; OTHER INFORMATION: Spacer peptide
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (26)..(34)
 ; OTHER INFORMATION: Amino acids 2-10 of the human GnRH hormone
 ; NAME/KEY: MOD_RES
 ; LOCATION: (34)..(34)
 ; OTHER INFORMATION: Amidated glycine or glycinamide
 US-09-848-834A-10

Query Match 4.6%; Score 112; DB 10; Length 34;
 Best Local Similarity 100.0%; Pred. No. 0.01;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 95 FNNFTVSWLRVPKVSASHLE 115
 :|||||
 Db 1 FNNFTVSWLRVPKVSASHLE 21

RESULT 14
 US-09-815-242-12209
 ; Sequence 12209, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; TITLE OF INVENTION: Prokaryotes
 ; FILE REFERENCE: ELITRA 011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848

Query Match	4.4%	Score 109;	DB 10;	Length 745;
Best Local Similarity	18.7%;	Pred. No. 1;		
Matches 100;	Conservative 78;	Mismatches 175;	Indels 182;	Gaps 26

RESULT 15
US-09-972-484-52
Sequence 52, Application US/09972484
Patent No. US20020137245.1

US20200212745A1
 GENERAL INFORMATION:
 APPLICANT: Miller, Timothy J.
 Kiepfer, Sharon
 Reed, Albert Paul
 Jones, Elaine V.
 TITLE OF INVENTION: Canine Coronavirus S Gene and Uses
 THEREFOR
 NUMBER OF SEQUENCES: 59
 CORRESPONDENCE ADDRESS:

SEQUENCE DESCRIPTION:	SEQ ID NO:
US-09-972-484-52	52

Query Match	4.4%	Score 108.5	DB 10	Length 1101
Best Local Similarity	18.1%	Pred. No. 1.8		
Matches	83	Conservative	75	Mismatches 145; Indels 155; Gaps 21
QY	32	STINLIDINNDIISDISGSSSVITTPDAQLVPGINGKAIHLVNNSESEVIVHKAMDIEY	91	
Db	8	ATVPSLMTGTGVLLEISICYNDTV-----	11	11
QY	92	N-----DMFNPFVYFWLRP-----KVSASHLEOYGTNEYSTISMKRHSLSI---	135	
Db	44	GVIDDPPRYCALXNGTALKYGLTPPSYKETAISKWGHFYINGVNFSTFPIDCISFNLIT	103	
QY	136	-----GSGMSVSLKG--NNLIMTLKDSAGEVRITP-----	171	
Db	104	TGDSGAFWTLIATYSYTDAIWOVENTA--IKKVTYCNSHINNIKKCSQLTANLQNGYPPVAS	161	
QY	172	--NAYLAKWFFITITNDRSSANLYINGVLMSGAEITGLGAIRED-----	221	
Db	162	SEWGLVNRKSVYLLLPSEFSHTSVNITID-----LGMKRSGYGQPIASTLSNITLTP	210	
QY	222	LDRCCNNNOYVSIDKFEIFCKALNPKFEKELKMSYLSITPLRDRWGNPLRDVTEYLLIPV	281	
Db	211	MQ--DNNTID-----VYCIRSNQ-----FSYIVHSTCKSSIM-----	239	
QY	282	ASSSKDVQLKNITDYMWL-----TNAPSYNGKILNIYYRRLYN-----	325	
Db	240	-----DDVFNSTCDTVLYATAVIAIKTGTCPEFSEDKLNNLT--TFNFKCISLDPVGNACAFDV	293	
QY	326	KRYTPNNNEIDSPKSDQFIKLKYVSNNNEIITVYPKDGNAPNMLDIRLRVG-----	378	
Db	294	AARTTINE--QVRS-----LVYIEEGDVIYVPSNSGLHIDS-VLHDSICDYNITYG	345	

Tue Dec 3 12:26:17 2002

us-09-816-467-2.rapb

Page 8

Oy 379 AFGIPLYKKMEAVKLDLKTYSVOLKLYDDKNASLGLV 416
| : : : : | : : : : | : : : :
Db 346 RTGVGIIROTNSTLLSGLYITSLSGDLGFKNVSDGVI 383

Search completed: December 1, 2002, 11:32:07
Job time : 24 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2002, 11:20:07 ; Search time 43 Seconds

(without alignments)
1035.122 Million cell updates/sec

Title: US-09-816-467-2

Sequence: 1 MWFSTPFPSTSKNLDKWD.....DKLGGDWYFPTDESGTND 463

Scoring table: BLOSUM62

Gapop 10.0 , Gapect 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

- 1: PIR_73:*
- 2: PIR1:*
- 3: PIR2:*
- 4: PIR3:*
- 5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	2451	99.8	1315	1 BRCLTN
2	653.5	26.6	1296	1 BRCLTN
3	644.5	26.2	1297	1 BRCLTN
4	643.5	26.2	1296	1 BRCLTN
5	617	25.1	1291	1 BRCLTN
6	590	24.0	1291	1 BRCLTN
7	578.5	23.6	1252	1 BRCLTN
8	571	23.2	1252	1 BRCLTN
9	560.5	22.8	1251	1 BRCLTN
10	517	21.1	1274	1 BRCLTN
11	430	17.5	1291	1 BRCLTN
12	430	17.5	1291	1 BRCLTN
13	411.5	16.8	1276	1 BRCLTN
14	406	16.5	1285	1 BRCLTN
15	269.5	11.0	1162	1 BRCLTN
16	263	10.7	1162	1 BRCLTN
17	244.5	10.0	1136	1 BRCLTN
18	244.5	10.0	1136	1 BRCLTN
19	244.5	10.0	1136	1 BRCLTN
20	217.5	8.9	1193	1 BRCLTN
21	217	8.8	1193	1 BRCLTN
22	214	8.7	1165	1 BRCLTN
23	212	8.6	1165	1 BRCLTN
24	201.5	8.2	1165	1 BRCLTN
25	143	5.8	143	1 BRCLTN
26	142	5.8	143	1 BRCLTN
27	136.5	5.6	142	1 BRCLTN
28	134.5	5.5	1225	1 BRCLTN
29	134.5	5.5	1225	1 BRCLTN

30	132	5.4	1127	2 T8317	ORF MSV156 hypthe
31	131.5	5.4	2925	2 T00133	RNA-directed RNA p
32	131	5.3	755	2 T41912	structural phospho
33	129	5.3	937	2 C97168	glycosyltransferas
34	129	5.3	1417	2 T18418	hypothetical prote
35	129	5.3	4436	2 E71086	hypothetical prote
36	128.5	5.2	1272	2 H82926	conserved hypotet
37	128	5.2	669	2 E71610	hypothetical prote
38	128	5.2	987	2 A64474	hypothetical prote
39	128	5.2	1225	2 A36607	hypothetical prote
40	128	5.2	1411	2 T18417	E2 glycoprotein
41	128	5.2	2523	2 T18477	hypothetical prote
42	127	5.2	1645	2 F98907	hypothetical prote
43	127	5.2	2391	2 T18410	phage-related prot
44	127	5.2	2496	2 A71616	cardamoyl-phosphat
45	126.5	5.2	1132	2 H82887	secreted protein p
					hypothetical prote
ALIGNMENTS					
RESULT 1					
BRCLTN					
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani					
N:Alternate names: tetanus neurotoxin					
C:Species: Clostridium tetani					
C>Date: 31-Mar-1988 #sequence-revision 31-Mar-1988 #text-change 03-Jun-2002					
C:Accession: A25689; A25757; B25194; B25194; A60759; S69348; S09364					
R:Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hude, E.					
EMBO J. 5, 2495-2502, 1986					
A:Title: Tetanus toxin: primary structure, expression in E. coli, and homology with b					
A:Reference number: A25689; MUID:87055814; PMID:3536478					
A:Accession: A25689					
A:Molecule type: DNA					
A:Residues: 1-1315 <EIS>					
A:Cross-references: GB:X06214; NID:940773; PIDN:CAA29564.1; PID:940770					
R:Fairweather, N.F.; Lyness, V.A.					
Nucleic Acids Res. 14, 7809-7812, 1986					
A:Title: The complete nucleotide sequence of tetanus toxin.					
A:Reference number: A25757; MUID:87040747; PMID:3774547					
A:Accession: A25757					
A:Molecule type: DNA					
A:Residues: 1-1315 <FAI>					
A:Cross-references: GB:X06214; NID:940773; PIDN:CAA29564.1; PID:940774					
R:Fairweather, N.F.; Lyness, V.A.					
J. Bacteriol. 165, 21-27, 1986					
A:Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C					
A:Reference number: A25194; MUID:86085672; PMID:3510187					
A:Accession: A25194					
A:Molecule type: DNA					
A:Residues: 743-1315 <FA2>					
A:Cross-references: GB:M12739; NID:9144920; PIDN:AAA23282.1; PID:9144921					
A:Accession: B25194					
A:Molecule type: protein					
A:Residues: 865-894 <FA3>					
R:Matsumoto, M.; Lei, D.-L.; Sugimoto, N.; Ozutsu, K.; Okabe, T.					
Infect. Immun. 57, 3588-3593, 1989					
A:Title: Isolation, purification, and characterization of fragment B, the NH-2-termin					
A:Reference number: A60759; MUID:9005436; PMID:2478476					
A:Accession: A60759					
A:Molecule type: protein					
A:Residues: 461-475 <MAT>					
R:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.					
A:Immunol. 142, 394-402, 1989					
A:Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.					
A:Reference number: J50096; MUID:89093918; PMID:2463305					
A:Contents: annotation; epitope region					
R:Schlauer, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; Dasgupta,					
Nature 359, 832-835, 1992					
A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo					
A:Reference number: S27125; MUID:93063293; PMID:1331807					
A:Contents: annotation					

R:de Filipis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
 Eur. J. Biochem. 229, 61-69, 1995
 A:Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
 A:Reference number: S69348; MUID:95262688; PMID:7744050
 A:Accession: S69348
 A:Molecule type: protein
 A:Residues: 2-31 <DEF>
 C:Comment: The source of this protein was an extrachromosomal plasmid.
 C:Comment: The precursor is cleaved by endogenous proteinase activity to form light (fragment A) and heavy (fragment B) chains. The amino end of the heavy chain (fragment B) binds to ganglioside GM1 on presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of synaptic vesicles.
 A:Description: blocks neuroexocytosis via hydrolysis of a Glu-Phe peptide bond in synaptobrevin.
 C:Superfamily: tetanus toxin
 C:Keywords: hydrolyase; metalloproteinase; neurotoxin; transmembrane protein; zinc
 F:2-457/Product: tetanolysin light chain (fragment A) #status predicted <TML>
 F:461-1315/Product: tetanolysin heavy chain (fragment B) #status experimental <TTM>
 F:461-864/DNA: channel forming (fragment B) #status predicted <TXS>
 F:865-1315/DNA: channel forming (fragment C) #status predicted <TXC>
 F:233-237/Binding site: zinc (His) #status predicted
 F:234/Active site: Glu #status predicted

Query Match 99.8%; Score 2451; DB 1; Length 1315;
 Best Local Similarity 100.0%; Pred. No. 6.3e-145; Indels 0; Gaps 0;
 Matches 462; Conservative 0; Mismatches 0;

QY 2 VSTPIPEFSYKSLDCVNEEDIDVILKSTILINDINNDIISDGSFNSVITPPDAQ 61
 DB 854 VSTPIPEFSYKSLDCVNEEDIDVILKSTILINDINNDIISDGSFNSVITPPDAQ 913
 QY 62 LVPGLNGKAIHLVNNSESEYVHKADIEYDNFNNFTYFSLRVKVSASHLEQYGTNE 121
 DB 914 LVPGLNGKAIHLVNNSESEYVHKADIEYDNFNNFTYFSLRVKVSASHLEQYGTNE 973
 QY 122 YSTIISMKKHSLSIGGWSYSLKGNLWTLKDSAGEVROITPRDLPDFNAVLAKKWF 181
 DB 974 YSTIISMKKHSLSIGGWSYSLKGNLWTLKDSAGEVROITPRDLPDFNAVLAKKWF 1033
 QY 182 ITTNDRLSSANLYINGVLMGSAETGLGAIREDNNITLKLDRNNNOYVSDKRIIFC 241
 DB 1034 ITTNDRLSSANLYINGVLMGSAETGLGAIREDNNITLKLDRNNNOYVSDKRIIFC 1093
 QY 242 KALNPKREIEKLYSYSTIFLRDPMGNPLRDYELIIPVASSSKDVLKNTIDYMLFN 301
 DB 1094 KALNPKREIEKLYSYSTIFLRDPMGNPLRDYELIIPVASSSKDVLKNTIDYMLFN 1153
 QY 302 APSYTGKELIYRRLYNGLFTIKRTPNNEIDSFYKSGDFIKLYSVNNNEHIVGYPK 361
 DB 1154 APSYTGKELIYRRLYNGLFTIKRTPNNEIDSFYKSGDFIKLYSVNNNEHIVGYPK 1213
 QY 362 DGAENFLDLIRLVGYAPGIPLYKKMEAVKLRDLKTYVQKLKYDKNASLGLVGHNG 421
 DB 1214 DGAENFLDLIRLVGYAPGIPLYKKMEAVKLRDLKTYVQKLKYDKNASLGLVGHNG 1273
 QY 422 QIGNDPMDILIASNMWFNHLKDKILGCDMIFVPTDEGWTND 463
 DB 1274 QIGNDPMDILIASNMWFNHLKDKILGCDMIFVPTDEGWTND 1315

RESULT 2

BTCIAB
 bontoxilysin (EC 3.4.24.69) A precursor - Clostridium botulinum

N:Alternate names: botulinum neurotoxin type A

C:Species: Clostridium botulinum

C:Date: 31-Mar-1993 #sequence, revision 31-Mar-1993 #text, change 18-Jun-1999

C:Accession: A35294; S09492; S68220; A33401; A53884; A60025; A27000

R:Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.

J. Biol. Chem. 265, 9153-9158, 1990

A:Title: The complete sequence of botulinum neurotoxin type A and comparison with other
 A:Reference number: A35294; MUID:90264400; PMID:2160960
 A:Accession: A35294

A:Molecule type: DNA
 A:Residues: 1-1296 <BIN>
 A:Cross-references: GB:M30196; NID:q144864; PIDN:AAA23262.1; PID:q144865
 A:Experimental source: Strain 62A, subtype A
 R:Thompson, D.E.; Brehm, J.R.; Oultram, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson, R.; Thompson, D.E.; Brehm, J.R.; Oultram, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson, R.
 Eur. J. Biochem. 189, 73-81, 1990
 A:Title: The complete amino acid sequence of the Clostridium botulinum type A neurotoxin.
 A:Reference number: S09492; MUID:90235864; PMID:2185020
 A:Accession: S09492

A:Molecule type: DNA

A:Residues: 1, 'Q', '3-26, 'V', '28-1296 <THO>

A:Cross-references: EMBL:X52066; NID:q40381; PIDN:CAA35289.1; PID:q40382

A:Experimental source: NCTC 2916

R:Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.

FEBS Lett. 376, 41-44, 1995

A:Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin compo

A:Reference number: S67988; MUID:96096783; PMID:8521962

A:Accession: S67988

A:Molecule type: preliminary

A:Status: preliminary

A:Residues: 1-12 <PUT>

A:Cross-references: EMBL:D67030; DDBJ:D50421; NID:q2160224

R:Betley, M.J.; Somers, E.; Dasgupta, B.R.

Biochem. Biophys. Res. Commun. 162, 1388-1395, 1989

A:Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-t

A:Reference number: A33401; MUID:89350959; PMID:2669749

A:Accession: A33401

A:Molecule type: DNA

A:Residues: 1-35 <RET>

A:Cross-references: GB:M27892; NID:q144880; PIDN:AAA23269.1; PID:q551776

R:Gienez, J.A.; Dasgupta, B.R.

J. Protein Chem. 12, 351-363, 1993

A:Title: Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72, 45, 42,

A:Reference number: A53884; MUID:94000342; PMID:8397793

A:Accession: A53884

A:Status: preliminary

A:Molecule type: protein

A:Residues: 867-880; 1148-1217, 'Y', 1219 <GIN>

A:Experimental source: strain Hall

A:Note: sequence extracted from NCBI backbone (NCBI:139159); sequence modified after

R:Dasgupta, B.R.; Dekleva, M.L.

Biochimie 72, 661-664, 1990

A:Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and a

A:Reference number: A60025; MUID:91120847; PMID:2126206

A:Accession: A60025

A:Molecule type: protein

A:Residues: 2-47 <DAS2>

R:Binz, T.; Blaszi, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.

J. Biol. Chem. 269, 1617-1620, 1994

A:Title: Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.

A:Reference number: A49708; MUID:94124495; PMID:8294407

A:Contents: annotation

C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic sy

C:Gene: atx; botA

C:Function: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal-assoc

C:Superfamily: tetanus toxin

C:Keywords: disulfide bond; hydrolyase; metalloproteinase; neurotoxin; transmembrane p

F:2-444/Product: bontoxilysin A light chain #status experimental <LGH>
 F:445-1296/Product: bontoxilysin A heavy chain #status experimental <HNY>
 F:223,227/Binding site: zinc (His) #status predicted
 F:224/Active site: Glu #status predicted

Query Match 26.6%; Score 653.5; DB 1; Length 1296;
 Best Local Similarity 33.1%; Pred. No. 8.6e-33;

Db 1183 NVVKNREY-----RLATNASOAGVEKILSALEIPDVGNLSQVVMK 1224
 QY 401 -----VOLKLYDDKNASLGLVGTGHNQIGNDPRDILIASNWFENHL--KDKILG 448
 Db 1225 SKDDGIRNKKMNLQDNGNDIGFIFGLHYD-----NLAIVASNMVNRQVGNASRRFG 1279
 QY 449 CDWYFVPTDEGM 460
 Db 1280 CSWEFIPVDGGM 1291

RESULT 5

140631
 non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum
 C:Species: Clostridium botulinum
 C:Date: 12-Aug-1996 #sequence:revision 12-Aug-1996 #text_change 16-Jul-1999
 C:Accession: 140631; S48103; S48104; S36015
 R:Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.
 R:Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.
 A:Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botulinum
 A:Reference number: 140631; MUID:94122659; PMID:7764370
 A:Accession: 140631
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1291 <RES>
 A:Cross-references: EMBL:X71343; NID:g296148; PIDN:CAA50482.1; PID:g296149
 R:Campbell, K.D.; Collins, M.D.; East, A.K.
 J: Clin. Microbiol. 31, 2255-2262, 1993
 A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id
 A:Reference number: S48103; MUID:94013372; PMID:8408542
 A:Accession: S48103
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 634-761, 'E', '763-841, 'M', '843, 'T', '845, 'N', '847-994 <CAM1>
 A:Cross-references: EMBL:X70814; NID:g407778; PIDN:CAA50145.1; PID:g407779
 A:Experimental source: non-proteolytic strain 2129B (Scott)
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
 A:Accession: S48104
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 634-843, 'T', '845, 'N', '847-994 <CAM2>
 A:Cross-references: EMBL:X70819; NID:g407780; PIDN:CAA50150.1; PID:g407781
 A:Experimental source: non-proteolytic strain Ekund 2B (Colworth 229)
 C:Comment: Botulinum neurotoxin type B in these strains may possess a capable catalytic S
 C:genetics:
 A:gene: bontC/b
 C:superfamily: tetanus toxin
 C:keywords: metalloprotein; neurotoxin; transmembrane protein; zinc
 F:2-441/Product: botulinum neurotoxin type B light chain #status predicted <LGHT>
 F:442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HVT>
 F:230-234/Binding site: zinc (His) #status predicted
 F:231/Active site: Glu #status predicted

Query Match 25.1%; Score 617; DB 2; Length 1291;
 Best Local Similarity 31.0%; Pred. No. 1.6e-30;
 Matches 149; Conservative 94; Mismatches 191; Indels 46; Gaps 14;

QY 5 TPIPEFSYKMLDQWDEEDIDVILK-----KSTILNIDINDISDISGENSESV 54
 Db 834 TIIPEFIS-----TYSIEILIKFKNYKSEILNITLILNRNMLIDLSGYAKV 885
 QY 55 ITPYDPAOLVINGK-AIHLVNESSEVIYKAMDEIYDAMFNNTVSEMLRPKVSASH 113
 Db 886 EYDVGVL-----NDKNQFKLTSSADSKIRVTONIIFNSMFLDFSVSWIRLIPKYNDD 941
 QY 114 LEQYGTGEYSITSMKKHSLISGSGVSLKGNLWTLKDSAGEVQLFR-DLPKFN 172
 Db 942 IQNYIHNEHYITLNCMKN-----SGMKISIRGNRIWTLIDJNCKTSVFEFYIREDIS 996
 QY 173 AYLANKVVFTITNDRLSSANLYINGVLMGSAETTGALREDNNTILKDRCNNNOYV 232
 Db 997 EYI-NRNFVYITNN-LDNKAIYINGTLESNMDIKDIGEIVNVEITFKLDGDVDRIOFI 1054

QY 233 SIDKRIEFCALNPKEIEKLTYSITSLIFELDFPQGNPLRYTEYLLIPVASSKDVOLKN 292
 Db 1055 WKYFSTFNTQLONSNKEIKYKIOSYSEYLLKDFPQGNPLMYKNEYMPANAKNSYIKLVK 1114
 QY 293 ITDYMYLTNAPSYTNGKLNITYYRRLYNGLKFIKRYDPNNEI-DSFVSGDFIKLYSYN 351
 Db 1115 DSSVGEILIRSKYNQNSYINRYNLICEKFLIRRESQSQINDIVKREYIHLDLVH 1174
 QY 352 NNE-HIVGPRDGNAPFNLDRLKRGYNAFGIPLYKKMEAVKLRDLKTYVOLKLYDDKN 410
 Db 1175 HEEMRVYAY---KYKDEEKLFTSLISDSNEFYKTIKRYEDQPSYSCOLLFKDDE 1230
 QY 411 AS--LGLVGTGTH---NGQIGNDPRDILIASNWFENHLKDK---ILGCDWYFVPTDEGMT 461
 Db 1231 STDIDIGLIGIRHFYSGLVRKKYKDYFCISKWYKLEVKRRPKYSNLGCMWCFIPRDEGMT 1290

RESULT 6

148940
 bontoxilysin (EC 3.4.24.69) B precursor - Clostridium botulinum
 N:Alternate names: botulinum neurotoxin type B (Bont/B)
 C:Species: Clostridium botulinum
 C:Date: 19-Dec-1993 #sequence:revision 18-Nov-1994 #text_change 18-Jun-1999
 C:Accession: A48940; S48105; S21575; A42871; S07155; S08562; S07128; S08574
 R:Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Breilm, J.K.; Atkinson, T.; Minton, N.P
 Appl. Environ. Microbiol. 58, 2345-2354, 1992
 A:Title: Molecular cloning of the Clostridium botulinum structural gene encoding the
 A:Reference number: A48940; MUID:92384550; PMID:1514783
 A:Accession: A48940
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1291 <WHB>
 A:Cross-references: GB:M81186; NID:g144734; PIDN:AAA23211.1; PID:g144735
 A:Experimental source: type B, Danish
 A:Note: sequence extracted from NCBI backbone (NCBIN:112080, NCBI:P:112081); this publ
 R:Campbell, K.D.; Collins, M.D.; East, A.K.
 J: Clin. Microbiol. 31, 2255-2262, 1993
 A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific
 A:Reference number: S48103; MUID:94013372; PMID:8408542
 A:Accession: S48105
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 634-994 <CAM>
 A:Cross-references: EMBL:X70817; NID:g407782; PIDN:CAA50148.1; PID:g407783
 A:Experimental source: proteolytic type B, strain NCTC 7273
 R:Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.
 submitted to the EMBL Data Library, April 1992
 A:Description: Partial amino acid sequence of botulinum neurotoxin type B and compari
 A:Reference number: S21575
 A:Accession: S21575
 A:Molecule type: DNA
 A:Residues: 36-217, 'G', '219-224, 'S', '226-246 <SZB>
 A:Cross-references: EMBL:211934; NID:g40383; PIDN:CAA77991.1; PID:g40384
 R:Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars
 J. Biol. Chem. 267, 14721-14729, 1992
 A:Title: Minimal essential domains specifying toxicity of the light chains of tetanus
 A:Reference number: A42871; MUID:92340509; PMID:1634516
 A:Accession: A42871
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-313, 'S', '315-451 <KUR>
 A:Experimental source: strain Okra
 A:Note: sequence extracted from NCBI backbone (NCBI:P:109365)
 R:Dasgupta, B.R.; Datta, A.
 Biochimie 70, 811-817, 1988
 A:Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity w
 A:Reference number: S07155; MUID:89000987; PMID:3139097
 A:Accession: S07155
 A:Molecule type: protein
 A:Residues: 2-29, 'M', '31-45 <DAS>
 A:Accession: S08562
 A:Molecule type: protein
 A:Residues: 442-463, 'R', '465-467 <DA2>

Query Match	24.0%;	Score 590;	DB 1;	Length 1291;
Best Local Similarity	31.7%;	Pred. No. 7.8e-29;		
Matches 140;	Conservative 89;	Mismatches 184;	Indels 28;	Gaps 12;

QY	34	ILINDINNDIISDISGNSSVITYYPDAOLVPGINGK-AIHLVNNSSSEVIYHKRADIEYN	92
Db	865	ILNLRYKDNLDLDSYGAKVEYDGYEL---NDKNOFKLITSANSKIRYQONONITIN	920
QY	93	DMFNFTYSFWLRLVRKVSASHLEQYGTNEYSIISMKRKHSLSIGSGVSLSKGNLIMTL	152
Db	921	SVFLDESVSEFWIRIPRYKNDGIONYIHNETYIINCMMKN---SGMKISIRGNRIIMTL	975
QY	153	KDSAGEVQITFR-DLPDKFNAYLANKWVFTITINDLSANLYINGVLMGSAETTGCA	211
Db	976	IDINGKTSVFEPEYNIREDISEYI-NRWEFVITINN-LNNAKIYINGLESNTDILKIDIRE	1033
QY	212	IREDNNTILKLDRCNNNNQYVSIIDKPRIFCKALNPKEIEKIYTSYLSTFLRDFGNPLR	271
Db	1034	VIANGELIFKLDGDLDRQFIIMKRYTIRNTELSQSNIIEEYKIQSYSEYIKDFEGNPLM	1093
QY	272	YDTEYYLIPVASSSSKDVOAKNITDYYMLTINAPSYTNGKLIYYRRLYNGLKEIIRKYPN	331
Db	1094	YKNEYMFNAGKRNKNSYIKLKRPDGPVEILTRSKYQNSKNTINYNDLVIGKEFTIRKNS	1153
QY	332	NEI-DSFYKSGFILL-YSYNNNEHIVGYPKGMAFNNIDRIIRVGNADGILPYKRM	389
Db	1154	QSIDNDIVRKEEYIYLDLFNLMQEMRWYTY---KYFKKEEEKFLAPISDSDEFYNTIQ	1209
QY	390	AVKLRDLKTSYVOLKLYDDKNAS--LGIWGT--NGQIGNDPRADILLASNWFNHLKD	444
Db	1210	IKKEYDEQPTYSQQLFKKDESDSTDEIGLGIHREYFESGIVFEYKDEYCSIKWYLKEVR	1269
QY	445	K----ILGCDWYFVPTDEGWT	461
Db	1270	KPYNLKILGCMNQFIPKDEGWT	1290

RESULT 7	
521178	botulinum neurotoxin type E precursor - Clostridium botulinum

	Query Match	23.6%	Score 578.5	DB 2	Length 1252
	Best Local Similarity	30.1%	Pred. No. 3.9e-28		
	Matches 148	Conservative 82	Mismatches 159	Indels 103	Gaps 16
QY	7	IPFSYKNDGCVNEEDIDYL-----KSTIIINDINDIIDISGFENSVI	55		
Db	823	IPFSLSYTD-----DKLIIYFNKFPRIKSSVLMNRYKKNKYDTSYDSHIN	873		
QY	56	TYPDAQLVPGINGKAIHLVNNESSEVIYHKAMDIEVDNENFTVSFWLVPKXASHLE	115		
Db	874	INGDVYKYP-TNKKQFGIYNDKRLSEVNISQNDYIIYDNRYKNEFSISWVRIPNDKRIYN	932		
QY	116	QYGTNEYSIISSMKHSLISIGSGWSVSLKGNLIMWLKDSAGEVROTFTR-DLPDKENAY	174		
Db	933	V--NNEETIINCMDNN---SGWKVYSLNHNELIIMLQDNAGINQKLAIFYGNANGISDY	986		

[illegible]

RESULT 8
S33411
botulinum neurotoxin type F - Clostridium baratii
C:Species: Clostridium baratii
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C:Accession: S33411, S31860
R:Thompson, D.E.; Hutton, R.A.; East, A.R.; Allaway, D.; Collins, M.D.; Richardson, P.T
FEMS Microbiol. Lett. 108, 175-182, 1993
A:Title: Nucleotide sequence of the gene coding for Clostridium baratii type F neurotoxin
A:Reference number: S33411; MUID:93252228; PMID:8486245
A:Accession: S33411
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1268 <THO>
A:Cross-references: EMBL:X68262; NID:949138; PIDB:CAA48329.1; PID:949139
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

[illegible]

```

Db      1108  VDGANEYQLADYVSISAVEKTIKLRIRISNINYSNQMIINDSLGDNCTMNEKT-----1220
QY      406   YDDKNASLGLGVTHNQIGNDENRDIILIASNWFYHFKDKIL--GCDWYFVPTDEGW 460
        :      :      :      :      :      :      :      :      :      :
Db      1221  --NNGNDIGLGFH--LNN-----LVASWYKKNIRNTRNNGCGFWSTISKEHGW 1266

```

RESULT 9

botulinum neurotoxin type E precursor - Clostridium butyricum
C:Species: Clostridium butyricum
C:Date: 30-Jun-1992 #sequence_revision 15-May-1998 #text_change 16-Jul-1999
C:Accession: JH0256; S16145
R:Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
C:Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter release.
A:Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum t
A:Reference number: JH0256; M01D:92181428; PMID:1543481
A:Accession: JH0256
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-27, 'E', 29-1251 <POU>
A:Cross-references: EMBL:X62088; NID:940379
A:Experimental source: strains ATCC 43181 and ATCC 43755
R:Fujii, N.; Kimura, K.; Yashiki, T.; Indoh, T.; Murakami, T.; Tsuzuki, K.; Yokosawa,
J. Gen. Microbiol. 137, 519-525, 1991
A:Title: Cloning of a DNA fragment encoding the 5'-terminus of the botulinum type E t
A:Reference number: S16145; M01D:91237316; PMID:2033376
A:Accession: S16145
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-229, 'M', 231-252 <FUJ>
A:Cross-references: EMBL:X53180; NID:940407; PIDN:CAA37321.1; PID:940408
A:Experimental source: strain BL6340
C:Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter release.
C:Comment: The heavy chain mediates the binding of toxin to cell receptors while the
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin
F:12-422/Product: botulinum neurotoxin type E light chain #status predicted <lig>
F:1423-1251/Product: botulinum neurotoxin type E heavy chain #status predicted <HBA>
F:412-426/Disulfide bonds: #status predicted

Query Match:	22.8%	Score 560.5;	DB 2;	Length 1251;
Best Local Similarity	29.3%	Pred. No. 5,2e-27;		
Matches 147;	Conservative 83;	Mismatches 164;	Indels 107;	Gaps 17;

QY	1	IWEFT---	PIPEYSKNDLCWVNEDIEDVIL-----	-KSTJLINDINDDIISD	46
		:	:	:	:
DB	814	MYIDLNNSIPKRLSSYTD-----	-DKLLSYFNKKFRKIKSSVLMNFKKDKKYD		864
		:	:	:	:
QY	47	ISGFNSVTTYDQAOLPQINKAIIHVNSESSEYIVHAKMIEVNDMEPNFTVSFWLHV			106
		:	:	:	:
DB	865	TSQYSDNININGDYKYP--TNKNQGIYNDKJSEVNISONDYIITYDNKTKNFSISFWARI			923
		:	:	:	:
QY	107	PKYSASHLEQYGTENXSIISSMKKSLSIGSGMSVSLKGNLITWLKDSAGEROIYTER			165
		:	:	:	:
DB	924	PVNDKNIVAV--NNETJIIINCMDRNN-----SCGMKVLNHEIITWFLDQSGINOKLAENY			977
		:	:	:	:
QY	166	DLDPKENAYLANKWFEITTTNDRLSSANLYINGVLMGSMETGLGAIRDNNTITKLDBC			225
		:	:	:	:
DB	978	GNANGISDXYI--KNMIVFTITTDRLGSGKIYINGNLLDKKSLNLGLGIVHSDNILEKIVNC			1036
		:	:	:	:
QY	226	NNNNOVSVIDKPIFECKALNPKEIEKLYSYLSITFLRDFWGNPLRYDTEYVLIPIYAS--			283
		:	:	:	:
DB	1037	-SYTKIGIRYFNIFPEKLEDETETQYLLNNPRAANILKQFWGYYLLDYKREYLLANLVKPN			1095
		:	:	:	:
QY	284	-----SKDVOAKNTDYMVLNAPSYNGKLNIIYRRLYNGLAKTIKRYTPNNEIDSF			337
		:	:	:	:
DB	1096	NFTNRRTDSTLSINNIRSTJILLAN-----RLYSIKYKIDRVNNSSTNDUL			1141
		:	:	:	:
QY	338	VKSGDFIKLYVSY-----NNNEHIYQPKDGAENFLDRIIRLVGNAPG			381
		:	:	:	:
DB	1142	VKAND--QYVINVASKTHLLPLYADATATNNKEKTIKISSGGRNFQV-----VYVNSVG			1194
		:	:	:	:

OY 382 IPIYKKMEAVKRLDQITTYVQLKLYDKNASLGIVGTHNGOIGDNDPNDILIASWYENH 441
 Db 1195 -----NCTMNFKNNNNNIGLIGF-----KADIVASTWYTYH 1227
 OY 442 LKDKI--LGCWYFVPTDEGW 460
 Db 1228 MROVTNSNGFFWNITSEBHW 1248

RESULT 10

140813
 neurotoxin type F - Clostridium botulinum
 C:Species: Clostridium botulinum
 C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999
 R:Accession: 140813; S48108
 R:Fast, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, D.E.
 FEMS Microbiol. Lett. 96, 225-230, 1992
 A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.
 A:Reference number: 140644
 A:Accession: 140813
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1274 <RES>
 A:Cross-references: GB:M2906; NID:g144866; PIDN:AAA23263.1; PID:g144867
 R:Campbell, K.D.; Collins, M.D.; East, A.K.
 J. Clin. Microbiol. 31, 2255-2262, 1993
 A:Title: Gene probes for identification of the botulinus neurotoxin gene and specific Ig
 A:Reference number: S48103; MUID:94013372; PMID:8408542
 A:Accession: S48108
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 634-1002 <CAM>
 A:Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789
 C:Superfamily: tetanus toxin
 C:Keywords: neurotoxin

Query Match 21.1%; Score 517; DB 2; Length 1274;
 Best Local Similarity 29.1%; Pred. No. 2.7e-24;
 Matches 131; Conservative 99; Mismatches 160; Indels 60; Gaps 17;

OY 29 LKSTLNDINDNDITSDISGENSVITPPDQVLDPQINGKA-----IHLVNESS 79
 Db 865 IKDSSILDRKREYENKRFIDISYGSNI-----SINGNVYISTNRNOFGIYNRLS 914
 OY 80 EYIVHAKMDIEYNDMNFNTVSFWLRVPRKVSASHLDQYGN EYITSSMKKHSLSIGSG 138
 Db 915 EYVIAQNDITINSRIQNSISFWRIPK---HYKPMNHNREYTIINCNGNNV---SG 966
 OY 139 WSVSL---KGNMLIWTLKDSAGEVROITER-DLPKFNAYLANKWVFITTTNRLSSANL 194
 Db 967 WKISLRVADCEIIMTLODTSQMKELIRYEELNLSNIY NKWIFVTITNNRLGNSRI 1025
 OY 195 YINGVLMSAETTGATIEDNNITLKLDRCNNOYVSDIKRFICKALNPKREIEKLYT 254
 Db 1026 YINGNMLIVKRSISNIGDHYSDNILEFKIIVGCDEET-YVGIREFKVPENLEDKTEIEFTLS 1084
 OY 255 SYLSTFLPDEFGNPLRKYTEYLLIPVASSSKDYOALKNTDWAYLTNAPSYTGNLITY 314
 Db 1085 NEPDPSILKNTYNGNLLNKKYTLLENLKRDKYITLNN-SGIININQOQRVTEGSAFLNLY 1142
 OY 315 RRLVGLKFLIIRKYP--NNEIDSEYKSGDFIKLYSYNNNEHIVGYPKDGNAFNNDRL 372
 Db 1143 -KLYGVEVYIIRKNGPIDISNDNFVRKNDL--AYI--NVDRGVEYRLXADYKSEKEKI 1197
 OY 373 LRVGYNAGGIPLYKKMEAVKRLDQITTYVQLKLYDKNASLGIVGTHNGOIGDNDPNDIL 432
 Db 1198 IRTS-----NLNDLSGQITVWDSIGNNCTMNFONNNGNSNIGLIGHSNN-----L 1242
 OY 433 IASWYENHKLKDKIL--GCDWYFVPTDEGW 460
 Db 1243 VASSWYNNIRNIRNITSSNGCFSSISKENGW 1272

RESULT 11

S46431
 botulinum neurotoxin C1 - Clostridium botulinum phage 1C (strain C 468)
 M:Alternate names: BONT/C1 protein
 C:Species: Clostridium botulinum phage 1C
 A:Variatey: Strain C 468
 C:Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 16-Jul-1999
 R:Accession: S46431; S49107
 R:Hauser, D.; Eklund, M.W.; Boquet, P.; Popoff, M.R.
 Mol. Gen. Genet. 243, 631-640, 1994
 A:Title: Organization of the botulinum neurotoxin C1 gene and its associated non-toxi
 A:Reference number: S46426; MUID:94301293; PMID:8028579
 A:Accession: S46431
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1291 <HAN>
 A:Cross-references: EMBL:X72793; NID:g516171; PIDN:CAA51313.1; PID:g516175
 A:Experimental source: strain C 468
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993
 C:Superfamily: tetanus toxin

Query Match 17.5%; Score 430; DB 2; Length 1291;
 Best Local Similarity 26.4%; Pred. No. 7.5e-19;
 Matches 135; Conservative 90; Mismatches 172; Indels 114; Gaps 21;

OY 3 FSTPIPFYSYKMLDQWVNDNEEDIDVI-----LKKSTLNDINDNDITSDISGENSVIT 56
 Db 840 FQNTIPE-----NIFSTNNSLKDIINEFNNINDSKILSLQNRKNTLVPTSGYNAEYSE 895
 OY 57 YPDAQIPL-----GINKAIIHLVNNESSEYIVHAKMDIEYNDMNFNTVSFWLRVPR 108
 Db 896 EBDVQIPLPFPPDFKLGSGE-----DRGKIVYQNEIVYVNSWYESPISFWIRINK 948
 OY 109 VSASHLEQYGTNEYSTISSMKKHSLSIGSGWSYSLGNNMLIWTLKDSAGVROITER-DL 167
 Db 949 -WVSNLP-----GYTIDSVKNN-----SGWSIGIISNLFVFLKQNEDESGSINFSYDI 997
 OY 168 PDKFNAYLANKWVFITTTNDRLSANLYINGVLMGSAEITGLCAIEDNNITLKLDR--- 224
 Db 998 SNNARGY--NKWFYLVTNMMGNNKRIYINGKLIIDIRYKELGINSKITTIEINKIPD 1055
 OY 225 -----CANNNOYVSDIKRFICKALNPKREIEKLYSYLSTPLRDFWGNPLRVDTEYLLI 279
 Db 1056 TGLTTSDDNINMMIRDFYIFAKELDKDINILFNSLQYTNVVKDYWGNDLRYNKEYYVW 1115
 OY 280 PVASSSKDYOALKNTDWAYLTNAPSYTGNKLNYYRLTN-----GLKFIKRYTTRNET 334
 Db 1116 NI-----DYL---NRMYIANSRQIVFNTRRNNNDNEGYKIIIRIGNTN- 1158
 OY 335 DSFKSGDFIKLYSYNNNEHIVGYPKDGNAFNNDRLIRVAGYNAPGIPLYKKMEAVK-- 392
 Db 1159 DTRVAGDILYEDMTINNKAY-----NLFMKNETYVADHSHSTEDYIAIGRQETKDI 1210
 OY 393 -----LRDLKTSYVOL--KLYDKNAS-LGLVGTGTHNGOIGDNDPNDILIASWY- 438
 Db 1211 NDNIITFQIPNNNTYVYASQIFKSNFNGENISGISIGYRFRRLGD-----WYR 1260
 OY 439 FNHLKDKI-----LGCWYFVPTDE 458
 Db 1261 HNYLVPTVYKQGNVYASLESTTHNGFVYVSE 1291

RESULT 12

A49777
 botulinum neurotoxin type C1 precursor - Clostridium botulinum phage (type C, strain
 C:Species: Clostridium botulinum phage
 C:Date: 10-Mar-1994 #sequence_revision 07-Apr-1994 #text_change 23-Mar-2001
 R:Accession: S11291; A35396; S22166; A49777
 R:Hauser, D.; Eklund, M.W.; Kurazono, H.; Binz, T.; Niemann, H.; Gill, D.M.; Boquet,
 Nucleic Acids Res. 18, 4924, 1990
 A:Title: Nucleotide sequence of Clostridium botulinum C1 neurotoxin.
 A:Reference number: S11291; MUID:90370487; PMID:2204031
 A:Accession: S11291

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-84, 'P', 86-1291 <HAU>
 A:Cross-references: EMBL:X53751; NID:g14905; PIDN:CAA37780.1; PID:g14906
 R:Kimura, K.; Fujii, N.; Tsuzuki, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Takeshi, K.
 Biochem. Biophys. Res. Commun. 171, 1304-1311, 1990
 A:Title: The complete nucleotide sequence of the gene coding for botulinum type C-1 toxin
 A:Reference number: A35396; MUID:91024998; PMID:2222445
 A:Accession: A35396
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-669, 'R', 671-1291 <TS1>
 R:Tsuzuki, K.; Kimura, K.; Fujii, N.; Yokosawa, N.; Oguma, K.
 submitted to the EMBL Data Library, December 1991
 A:Description: Nucleotide sequence of the gene for one of the components of hemagglutinin
 A:Reference number: S22163
 A:Accession: S22166
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1291 <TS2>
 A:Cross-references: EMBL:X62389; NID:g558175; PIDN:CAA44263.1; PID:g40390
 R:Kimura, K.; Fujii, N.; Tsuzuki, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Oguma, K.
 Appl. Environ. Microbiol. 57, 1168-1172, 1991
 A:Title: Cloning of the structural gene for Clostridium botulinum type C-1 toxin and who
 A:Reference number: A49777; MUID:91282468; PMID:2059039
 A:Accession: A49777
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-607 <TS3>
 A:Cross-references: GB:D90210
 C:Superfamily: tetanus toxin
 C:Keywords: neurotoxin

Query Match 17.5%; Score 430; DB 2; Length 1291;
 Best Local Similarity 26.4%; Pred. No. 7.5e-19;
 Matches 135; Conservative 90; Mismatches 172; Indels 114; Gaps 21;

3 FSTPIPFYSKNDLCWVNEEDIDVI-----LKKSTILNDINDNDIISDSCFNSSVIT 56
 DB 840 FQMTIP-----NIFSTYNNSLKDIINEYNNINDSKILSLONKKNLVDTSQNAEYSE 895
 OY 57 YPDQOLVP-----GINKAITHLVNNESEVIVKAMDIENDMENFTVSFWLRVPK 108
 DB 896 EGDVQINPIPFDFKLGSSGE-----DRGKVIIVQENINIVNSMSEFSISFWIRNK 948
 OY 109 VSAHLEQYGTNEYSTISSMKKHSLSIGSGWSVSLKGNLIMTLKDSAGEVROITFR-DL 167
 DB 949 -WVSNNLP-----GYTIIDSVKNN-----SGWSIGIISNFIIVFTLKQEDSEOSINFSYDI 997
 OY 168 PDKENAYLANKWVEITITRLSSANLYINGVLMSAEITGLAIREDNNTITLKR-- 224
 DB 998 SNNAPGT--NKWFFVYITNNMGMKIKYINGKLIDITKVKELTGINSKTTTFELNKLPD 1055
 OY 225 -----CNNNNQVYSIDKFRIFCKALNKEIEKLYTSYSTIFELDFWGNPLRYDTEYLI 279
 DB 1056 TGLITSDSDNINMIRDFYIFAKELDGKDINILFNSLOQYTVNVRDYMGNDRYKNKEYYAV 1115
 OY 280 PVASSKSDVQKNTDQWYLTNAPSYNGKINITYRRLYN-----GKFIIRKRYTPNNEI 334
 DB 1116 NI-----DYL-----NRYMANSROIIVNTRRNNDPNEGKIIIRKIRGNIN- 1158
 OY 335 DSFFKSGDFIKLYSVNNNEHIVGPRDGNANFNLDRLIRGVANAPGIPLKKMEAVR-- 392
 DB 1159 DTRVRGDDILYFQMTINNKAV-----NLFMKNETPVADNNHSTEDYIATGLRQTKDI 1210
 OY 393 -----LRDLKTSYVOL--KLYDDKNAAS-LGLVGTNAGQGNPNRDLILASNNY- 438
 DB 1211 NDNIIFQIOPNNNTYVYASQIFKSNFNGENISGICSTGTTFYRFLGGD-----WYR 1260
 OY 439 FNHLKDKI-----LGCWVFEVPTDE 458
 DB 1261 HNYLVPTVYKQGNVYASLLESTSTHNGFVYVSE 1291

RESULT 13
 S11455
 botulinum neurotoxin type D - Clostridium botulinum
 C:Species: Clostridium botulinum
 C:Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
 A:Accession: S11455
 R:Binz, T.; Kurazono, H.; Popoff, M.R.; Eklund, M.W.; Sakaguchi, G.; Kozaki, S.; Krie
 Nucleic Acids Res. 18, 5556, 1990
 A:Title: Nucleotide sequence of the gene encoding Clostridium botulinum neurotoxin ty
 A:Reference number: S11455; MUID:91016853; PMID:2216736
 A:Accession: S11455
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1276 <BIN>
 A:Cross-references: EMBL:X54254; NID:g40395; PIDN:CAA38175.1; PID:g40396
 C:Superfamily: tetanus toxin
 C:Keywords: neurotoxin

Query Match 16.8%; Score 411.5; DB 2; Length 1276;
 Best Local Similarity 24.9%; Pred. No. 1.1e-17;
 Matches 127; Conservative 96; Mismatches 164; Indels 123; Gaps 19;

3 FSTPIPFYSKNDLCWVNEEDIDVI-----LKKSTILNDINDNDIISDSCFNSSVIT 56
 DB 836 FENTMPF-----NIFSTYNNSLKDIINEYFNSINDSKILSLONKKNALVDTSQNAEVRV 891
 OY 57 YPDQOLVGINCKATHL-----VNNSESEVIVKAMDIENDMENFTVSFWLRVPK-VSA 111
 DB 892 GDNVOL-----NTIYNDNFKLSSGDKITIVNANNILYSAIYENSVSFWIKISKDTIN 945
 OY 112 SHLEQYGTNEYSTISSMKKHSLSIGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPDKF 171
 DB 946 SH-----NEITIINSIEGN-----SGMKLCIRNGINEMILQOVNRKYSKLIDYSESL 994
 OY 172 NAYLANKWVEITITNDRSSANLYINGVLMSAEITGLAIREDNNTITLKDRCNNNOY 231
 DB 995 HTGYTNKWFYVITNNIMYKLYINGELKQSKIDLEDEKLDKTIYEGIDENIDENOM 1054
 OY 232 VSIDKFRIFCKALNKEIEKLYTSYSTIFELDFWGNPLRYDTEYLIIPVASSKDVOLK 291
 DB 1055 LMRPNTFISKLSMEDINIVLEGQILRNVIKDYKWNPLKFDIYEYIIN----- 1103
 OY 292 NITDQWYLTNAPSYNGKINITY--RRLYNGLEKFIKRYTPNNEIDSEVKSDFIKLYV 348
 DB 1104 --DNYIDRYIADE-SNVLIVQYPPDRSKLYTGNPITIKSVSDKNPY--SRLINDGNILHM 1159
 OY 349 STNNNEHIVGPRDGNANFNLDRLIRGVANAPGIPLKKMEAVKLRDLPKT-YSVQ----- 402
 DB 1160 LYSNRKYM-----IRDPITIIATOGGEC 1184
 OY 403 -----LKLYPD-KNASLGLVGTN-----GOIGDENRDLIASN--WYFN-- 440
 DB 1185 QNCVYALKQSNLNGNIGIIFISIKNTVSKNKKYCSQIFSSSRRENTMLADYIKPWFSEKN 1244
 OY 441 -----HLKDKIIGCD--WYFVPTDEG 460
 DB 1245 AYPVAVTYVETKILSTSSFWKFIISRDPGW 1274

RESULT 14
 S70582
 botulinum neurotoxin type Dsa precursor - Clostridium botulinum phage d-sa
 C:Species: Clostridium botulinum phage d-sa
 A:Note: host Clostridium botulinum type D (strain South Africa)
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Jun-2000
 A:Accession: S70582
 R:Moriishi, K.; Koura, M.; Abe, N.; Fujii, N.; Fujinaga, Y.; Inoue, K.; Ogumad, K.
 Biochim. Biophys. Acta 1307, 123-126, 1996
 A:Title: Mosaic structures of neurotoxins produced from Clostridium botulinum types C
 A:Reference number: S70582; MUID:96283801; PMID:8679691
 A:Accession: S70582
 A:Status: nucleic acid sequence not shown

A: Molecule type: DNA
 A: Residues: 1-1285 <MOR>
 A: Cross-references: EMBL:D38442; NID:q1374775; PIRN:BAA07477.1; PID:q1374776
 C: Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit neu-
 C: disulfide bond. The heavy chain mediates the binding of toxin to the presynaptic mem-
 C: Superfamily: tetanus toxin
 C: Keywords: disulfide bond; neurotoxin; transmembrane protein
 F: 1-447/Product: botulinum neurotoxin type Dsa light chain #status predicted <MAT1>
 F: 448-1285/Product: botulinum neurotoxin type Dsa heavy chain #status predicted <MAT2>

Query Match 16.5% Score 406; DB 2; Length 1285;
 Best Local Similarity 26.8%; Pred. No. 2.3e-17;

Matches 133; Conservative 87; Mismatches 190; Indels 86; Gaps 19;

```

QY 3 ESTPIPEFSYKMDGVNEDIDVI-----LKKSTILNDINNDIISDGFSSVIT 56
Db 836 FENTIPF-----NIFSYTNLSLKDMINEFNSINDSKILLQKKRTIADTSYNEVRV 891
QY 57 YPDAGLV-----GNGKAIHLVNNSESEYVHKAMIDYNDMNFETVSEWLRVPR 108
Db 892 EGNVOLNPIPEDFELGSSG-----DDRGKITVQNMENIYVAMAYESFISFWIRNK 944
QY 109 VSASHLEQYGTNEYSIISSMKHSISGSGVSLKGNLITWTLKDSAGEVRQITPR-DL 167
Db 945 -WVSULP-----GYTIIISYKNN-----SGWSIGTISNPLVETLKONENSEQDINFSDYI 993
QY 168 PDKRPAIYANKWFTTTTNDRLSSANLYINGVLMGSAETGLGAIREDNNITLKDRCNN 227
Db 994 SKNAGY--NKFVETITTNMGMNIYINGKLIIDTIKVELTGINFSTIIFQMKIRN 1051
QY 228 NNGYVS-----IDKRIECKALNPKIEIKLYTSYSTLFLRDFWGNPLKRIDEYILI 279
Db 1052 TGLTSDSDNINMMIRDFYIAKELDDKDINLENSIQYTNVVKYMGNDLYRDEKYMI 1111
QY 280 PVASSKDVQLKNTIDYMYLTNAPSYTNGKLNITYRRLYNGLKFIKRTYTPNNEIDSEVK 339
Db 1112 NVNMYNMYMSKKG-----NGIYFNRKNNNDPN---EGYKIIIRKRGTFN-DTRVK 1159
QY 340 SGDEFIKLYVSNNNEHIYV-PRKDNAPFNNDRLIRVGYNAPGIPLYKKMEAVKRLDKT 398
Db 1160 GENLYVNTTIDNKQYSLGMMKPSRNLGTDLVPLGALDQPMDEIRKYGSEFIQPCNTFDY 1219
QY 399 YSVOLKLYDDKNAS-IGD--VGTHNGOIGDNPNDIILASNMVFNH-----LKD----- 444
Db 1220 YASQLFSSNATNTRLGLISGYSYFKLGD-----YWFNHEYLLIPYIKTEHYAS 1269
QY 445 --KILGDMWFVPTDE 458
Db 1270 LLESTSTHWVFPASE 1285

```

RESULT 15
 I40817

botulinum toxin nontoxic component - Clostridium butyricum

C: Species: Clostridium butyricum
 C: Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Oct-1999

C: Accession: I40817
 R: Fujii, N.; Yokosawa, N.; Oguma, K.; Yashiki, T.; Takeshi, K.; Isogai, E.;
 Microbiol. Immunol. 37, 395-398, 1993

A: Title: Similarity in nucleotide sequence of the gene encoding nontoxic component of bc
 Mashike.

A: Reference number: I40817; MUID:93360835; PMID:8355622

A: Accession: I40817
 A: Status: preliminary; translated from GB/EMBL/DBD

A: Molecule type: DNA

A: Residues: 1-1162 <RES>
 A: Cross-references: GB:012739; NID:9285709; PIDN:BAA02231.1; PID:g285710

Query Match 11.0%; Score 269.5; DB 2; Length 1162;
 Best Local Similarity 25.6%; Pred. No. 6.7e-09;
 Matches 114; Conservative 67; Mismatches 144; Indels 121; Gaps 21;
 QY 15 LDCWVNDNEDIVITLKSTIILMDI----- 39

```

Db 749 LKQTNINTEKSHLLMONSFSLDFDLDIQNMKNLFNLYTELLIKQTSPELSLAFQ 808
QY 40 --NNDIISDTSGFSSSVITYP-DAQLYPGINKAIHLVNNSESEYVHKAMIDYNDMFN 96
Db 809 EODNNVIEDTSCKN-TLVEYFDIGLYGGINNAHILTCANON--IKFTNDFEENGLTN 864
QY 97 NFTVSEWLRVPRVSHLEQYGTNEYSIISSMKHSISGSGVSLKGNLITWTLKDSAGEVRQITPR-DL 151
Db 865 NESTYFWLR-----NLNQTIRSKL-----IGSKEDNCGWELIPENNGLVFN 906
QY 152 LKDSAGEVRQITPRDLDPKFNAYLAN---KVFYTTTNDRLSSANL-YINGVLMGSAEI 206
Db 907 IIDSNGEK-----NITLSINISKNWYHLYISINRLKDOLLIFIDNILVANEDI 955
QY 207 TGLGAIREDNNITLKDRCNNNOYVSIDKPRIFCKALNPKIEIKLYTSYSTLFLRDFW 266
Db 956 KEILNIVSSDIISLSD--NNVY--IEGLSVLAKTINSNELLDYFSDLNNSYIRNF 1010
QY 267 GNPRLRYDEY---YLIFVASSKDVQLKNTIDYMYLTNAPSYTNGKLNITYRRLYNGIK 322
Db 1011 EETIQNRYTELEFNYPPEIANKIEQNNNI---YLSN--NNENSINL-----FKPLK 1057
QY 323 FIIRKTYPN-----NEIDSFVKSGDEFIKLYVSYNNE-HIVGPRKQGNAP--NN--L 369
Db 1058 FKILNTPNNOYVQKMDVEIFSVLDGTEKYLDISIDNRRIQLVNKNNAKYFTIINDIPI 1117
QY 370 DRILRVGYNAPGIPLYKKMEAVKRLD 395
Db 1118 SNCLTLYNNVNYVL-----SIKMOD 1138

```

Search completed: December 1, 2002, 11:24:52
 Job time : 54 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2002, 10:23:47 ; Search time 25 Seconds

(without alignments)
768.142 Million cell updates/sec

Title: US-09-816-467-2

Sequence: 1 MVEFTPIPFYSKINDCWVD.....DKILGDMYFVPTDEGTFND 463

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2451	99.8	1314	1	TETX_CLOTE
2	653.5	26.6	1295	1	BXA1_CLOBO
3	644.5	26.2	1296	1	BXG_CLOBO
4	643.5	26.2	1295	1	BXA2_CLOBO
5	590	24.0	1290	1	BXB_CLOBO
6	565	23.0	1250	1	BXE_CLOBO
7	560.5	22.8	1250	1	BXE_CLOBO
8	517	21.1	1274	1	BXE_CLOBO
9	430	17.5	1290	1	BXC1_CLOBO
10	411.5	16.8	1276	1	BXD_CLOBO
11	269.5	11.0	1162	1	BXEN_CLOBO
12	263	10.7	1162	1	BXEN_CLOBO
13	244.5	10.0	1196	1	BXEN_CLOBO
14	134.5	5.5	1225	1	BXCN_CLOBO
15	131	5.3	755	1	VG12_CVPR8
16	128	5.2	987	1	P100_HSV7J
17	128	5.2	1225	1	VG12_CVPR1
18	125	5.1	1449	1	VG12_CVPR1
19	123	5.0	1447	1	VG12_CVPR1
20	122.5	5.0	1447	1	VG12_CVPR1
21	119.5	4.9	2366	1	TOXB_CLODI
22	119	4.8	451	1	V110_FOPPV
23	119	4.8	1447	1	VG12_CVPR1
24	119	4.8	1447	1	VG12_CVPR1
25	118.5	4.8	1279	1	APU_THESA
26	117	4.8	1447	1	APU_THESA
27	117	4.8	1447	1	APU_THESA
28	116.5	4.7	634	1	YCX3_EUGGR
29	116	4.7	1138	1	C7AA_BACTU
30	116	4.7	1177	1	Y307_MYCGE
31	115.5	4.7	839	1	YDBB_HAEIN
32	115	4.7	993	1	NISB_LACLA
33	114	4.6	1729	1	RRP5_YEAST

34	113.5	4.6	613	1	CGAA_CLOBI	Q45882 Clostridium
35	113.5	4.6	967	1	Y560_CLOAB	P33747 Clostridium
36	113.5	4.6	1127	1	MDM1_YEAST	Q01846 saccharomyc
37	113	4.6	1138	1	C7AB_BACUR	Q45708 bacillus th
38	113	4.6	1449	1	VG12_CVPR5	P18450 porcine tra
39	112.5	4.6	1024	1	RPOB_PLAFA	P21421 plasmodium
40	112	4.6	806	1	SYFR_MYNGE	P47437 mycoplasma
41	112	4.6	849	1	PHSG_SYNGE	P73511 synecocyst
42	112	4.6	1386	1	RPOD_MARPO	P06274 marchantia
43	112	4.6	1956	1	ATX1_PLAFA	Q04956 plasmodium
44	112	4.6	3144	1	VP13_YEAST	Q07878 saccharomyc
45	111.5	4.5	592	1	Y036_HAEIN	Q57335 haemophilus

ALIGNMENTS

RESULT 1
TETX_CLOTE
ID TETX_CLOTE STANDARD: PRT: 1314 AA.
AC P04958:
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylysin).
OS Clostridium tetani.
OG Plasmid.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_Taxid=1513;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87053814; PubMed=3536478;
RA Etsel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA Weller U., Hudel M., Habermann E., Niemann H.;
RT "Tetanus toxin: primary structure, expression in E. coli, and
RT homology with botulinum toxins";
RL EMBO J. 5:2495-2502(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CM3911;
RX MEDLINE=87040747; PubMed=3774547;
RA Fairweather N.F., Lyness V.A.;
RT "The complete nucleotide sequence of tetanus toxin.";
RL Nucleic Acids Res. 14:7809-7812(1986).
RN [3]
RP SEQUENCE OF 742-1314 FROM N.A.
RX MEDLINE=86085672; PubMed=3510187;
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
RT fragment C in Escherichia coli.";
RL J. Bacteriol. 165:21-27(1986).
RN [4]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=90201034; PubMed=2108021;
RA Krieglstein K., Henschen A., Weller U., Habermann E.;
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups
RT in tetanus toxin.";
RL Eur. J. Biochem. 168:39-45(1990).
RN [5]
RP PARTIAL SEQUENCE.
RX MEDLINE=92037649; PubMed=1935979;
RA Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;
RT "Limited proteolysis of tetanus toxin. Relation to activity and
RT identification of cleavage sites.";
RL Eur. J. Biochem. 202:41-51(1991).
RN [6]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93010948; PubMed=1396558;
RA Schiavo G., Poullain B., Rossetto O., Benfenati F., Tauc L.,
RA Montecucco C.;
RT "Tetanus toxin is a zinc protein and its inhibition of
RT neurotransmitter release and protease activity depend on zinc.";

EMBO J. 11:3577-3583(1992).
 [7]
 IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=93063293; PubMed=1331807;
 RA Sciarro G., Benfenati F., Poulin B., Rossetto O., de Laureto P.P.,
 Dasgupta B.R., Montecucco C.,
 "Tetanus and botulinum B neurotoxins block neurotransmitter release
 by proteolytic cleavage of synaptobrevin.";
 RL Nature 359:832-835(1992).
 [8]
 X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
 RX MEDLINE=97475217; PubMed=9334741;
 RA Umhland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
 Sax M.;
 "Structure of the receptor binding fragment HC of tetanus
 neurotoxin.";
 RL Nat. Struct. Biol. 4:788-792(1997).
 CC -I- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-1-PHE-77
 BOND OF SYNAPTOSOMAL-2.
 CC -I- CATALYTIC ACTIVITY: HYDROLYSIS OF 76-GLN-1-PHE-77 BOND IN
 SYNAPTOSOMAL-2.
 CC -I- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
 YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
 AND ARE NON-TOXIC AFTER SEPARATION.
 CC -I- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
 GANGLIOSIDE RECEPTORS.
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
 CC or send an email to license@isb-slb.ch).

 CC EMBL: X0436; CAA28033.1; -;
 DR EMBL: M12739; AAA23282.1; -;
 DR EMBL: X06214; CAA29564.1; -;
 DR PIR: A25689; BTCLTN.
 DR PDB: 1AF9; 29-APR-98.
 DR PDB: 1ABD; 14-OCT-98.
 DR MEROPS: M27.001; -;
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_Mpeptidase.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOTOXILYSIN.
 DR PRODOM: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
 KW 3D-structure.
 KM
 FT INIT MET 0 0
 FT CHAIN 1 456 TETANUS TOXIN LIGHT CHAIN.
 FT CHAIN 457 1314 TETANUS TOXIN HEAVY CHAIN.
 FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 233 233 BY SIMILARITY.
 FT METAL 236 236 ZINC (CATALYTIC) (BY SIMILARITY).
 FT TRANSMEM 226 246 POTENTIAL.
 FT TRANSMEM 669 689 POTENTIAL.
 FT DISULFID 438 466 INTERCHAIN.
 FT DISULFID 1076 1092
 FT SEQUENCE 1314 AA; 150550 MW; 134C3657133EF81D CRC64;
 Query Match 99.8%; Score 2451; DB 1; Length 1314;
 Best local similarity 100.0%; Pred. No. 4e-15;
 Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 VSTPIPFSTSKNLDKWDVNEEDIDVILKSTILNLDINNDIISDLSGFSNSVITTPDAQ 61

|||||
 Db 653 VFTPIPFSTSKNLDKWDVNEEDIDVILKSTILNLDINNDIISDLSGFSNSVITTPDAQ 912
 QY 62 LVPGINGKAHLVNNSESEVYVHKAMDIEYNDMNFVSWLPVPKVASHLBOYGTNE 121
 Db 913 LVPGINGKAHLVNNSESEVYVHKAMDIEYNDMNFVSWLPVPKVASHLBOYGTNE 972
 QY 122 YSIISMKKHSLSIGSGSVSLKGNLITWTLKDSAGEVROITFRDLPKFNAYLANKWVF 181
 Db 973 YSIISMKKHSLSIGSGSVSLKGNLITWTLKDSAGEVROITFRDLPKFNAYLANKWVF 1032
 QY 182 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDDNITLTKDRCNNNOYVSIKFRIFC 241
 Db 1033 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDDNITLTKDRCNNNOYVSIKFRIFC 1092
 QY 242 KALNKEIEKLYTSLSTFTFLRDFWGNPLRDTVEYLLIPVASSKDVOLKNTDPMYLTN 301
 Db 1093 KALNKEIEKLYTSLSTFTFLRDFWGNPLRDTVEYLLIPVASSKDVOLKNTDPMYLTN 1152
 QY 302 APSYNGKLNITYRRLYNGQLKFIIRKYPNNEIDSFVKSQDFIKLYSYNNEHIVGYPK 361
 Db 1153 APSYNGKLNITYRRLYNGQLKFIIRKYPNNEIDSFVKSQDFIKLYSYNNEHIVGYPK 1212
 QY 362 DGNAPNNDRLIRVGNAPGTPLYKMEAVKRLDKTSVQLKLYDDKNASGLVGTHTG 421
 Db 1213 DGNAPNNDRLIRVGNAPGTPLYKMEAVKRLDKTSVQLKLYDDKNASGLVGTHTG 1272
 QY 422 QIGNDPNRDILIASNMWYFNHDKILGCDWYFVPTDEGTND 463
 Db 1273 QIGNDPNRDILIASNMWYFNHDKILGCDWYFVPTDEGTND 1314
 RESULT 2
 BXAL_CLOBO STANDARD: PRT: 1295 AA.
 ID BXAL_CLOBO
 AC P10845; P18639; P01561;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BONT/A)
 DE (Bontoxilysin A) (BOTOX) [contains: Botulinum neurotoxin A, light-
 chain; Botulinum neurotoxin A, heavy-chain].
 GN BOTA OR BNA OR ATX.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 CC Clostridium.
 OX NCBL_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCCTC 2916;
 RX MEDLINE=90235864; PubMed=2185020;
 RA Thompson D.E., Brehm J.K., Oultram J.D., Swinfield T.-J.,
 J. Thompson G.C., Atkinson T., Melling J., Minton N.P.;
 "The complete amino acid sequence of the Clostridium botulinum type A
 neurotoxin, deduced by nucleotide sequence analysis of the encoding
 gene.";
 RT Eur. J. Biochem. 189:73-81(1990).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=62A;
 RX MEDLINE=90264400; PubMed=2160960;
 RA Binz B., Kiarzono H., Wille M., Frevent J., Wernars K., Niemann H.;
 "The complete sequence of botulinum neurotoxin type A and comparison
 with other clostridial neurotoxins.";
 RT J. Biol. Chem. 265:9153-9158(1990).
 RL [3]
 RP SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN=62A;
 RX MEDLINE=97016817; PubMed=8863443;
 RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
 "Organization and phylogenetic interrelationships of genes encoding
 components of the botulinum toxin complex in proteolytic Clostridium
 botulinum types A, B, and F: evidence of chimeric sequences in the

RT gene encoding the nontoxic nonhemagglutinin component.";
 RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
 RN [4]
 RP SEQUENCE OF 1-34 FROM N.A.
 RC STRAIN-Hall;
 RX MEDLINE=89350959; PubMed=2669749;
 RA Betley M.J., Somers E., Dasgupta B.R.;
 RT "Characterization of botulinum type A neurotoxin gene: delineation of
 the N-terminal encoding region.";
 RL Biochem. Biophys. Res. Commun. 162:1388-1395(1989).
 RN [5]
 RP SEQUENCE OF 1-18 FROM N.A.
 RC STRAIN-type A NIH;
 RX MEDLINE=96096783; PubMed=8521962;
 RA Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;
 RT "Molecular characterization of two forms of nontoxic-nonhemagglutinin
 components of Clostridium botulinum type A progenitor toxins.";
 RN FEBS Lett. 376:41-44(1995).
 RL [6]
 RP SEQUENCE OF 1-16.
 RX MEDLINE=84178501; PubMed=6370252;
 RA Schmidt J.J., Sartymoorthy V., Dasgupta B.R.;
 RT "Partial amino acid sequence of the heavy and light chains of
 botulinum neurotoxin type A.";
 RL Biochem. Biophys. Res. Commun. 119:900-904(1984).
 RN [7]
 RP SEQUENCE OF 1-46.
 RX Dasgupta B.R., Foley J., Niece R.;
 RT "Partial sequence of the light chain of botulinum neurotoxin type A.";
 RL Biochemistry 26:4162-4162(1987).
 RN [8]
 RP SEQUENCE OF 1-5 AND 44-456.
 RX MEDLINE=91120847; PubMed=2126206;
 RA Dasgupta B.R., Dekleva M.L.;
 RT "Botulinum neurotoxin type A: sequence of amino acids at the
 N-terminus and around the nicking site.";
 RL Biochimie 72:661-664(1990).
 RN [9]
 RP SEQUENCE OF 448-464 AND 872-895.
 RX MEDLINE=89024662; PubMed=3178218;
 RA Sathymoorthy V., Dasgupta B.R., Foley J., Niece R.L.;
 RT "Botulinum neurotoxin type A: cleavage of the heavy chain into two
 halves and their partial sequences.";
 RL Arch. Biochem. Biophys. 266:142-151(1988).
 RN [10]
 RP SEQUENCE OF 448-482.
 RX MEDLINE=85285016; PubMed=3896784;
 RA Shone C.C., Hambleton P., Melling J.;
 RT "Inactivation of Clostridium botulinum type A neurotoxin by trypsin
 and purification of two tryptic fragments. Proteolytic action near
 the COOH-terminus of the heavy subunit destroys toxin-binding
 activity.";
 RL Eur. J. Biochem. 151:75-82(1985).
 RN [11]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94063091; PubMed=824676;
 RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,
 RA Benfenati F., Wilson M.C., Montecucco C.;
 RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
 COOH-terminal peptide bonds.";
 RL FEBS Lett. 335:99-103(1993).
 RN [12]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94124495; PubMed=8294407;
 RA Blinz T., Blas J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
 RA Jahn R., Niemann H.;
 RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
 RL J. Biol. Chem. 269:1617-1620(1994).
 RN [13]
 RP MUTAGENESIS OF GLU-261, PHE-265 AND TYR-365.
 RX MEDLINE=21556941; PubMed=11700044;
 RA Rigoni M., Caccin P., Johnson E.A., Montecucco C., Rossetto O.;
 RT "Site-directed mutagenesis identifies active-site residues of the

RT light chain of botulinum neurotoxin type a.";
 RL Biochem. Biophys. Res. Commun. 288:1231-1237(2001).
 RN [14]
 RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
 RX MEDLINE=98455071; PubMed=9783750;
 RA Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.;
 RT "Crystal structure of botulinum neurotoxin type A and implications
 for toxicity.";
 RL Nat. Struct. Biol. 5:898-902(1998).
 CC -1- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
 binds with high affinity to peripheral neuronal presynaptic
 membrane, is then internalized by receptor-mediated endocytosis.
 CC The C-terminus of the heavy chain (H) is responsible for the
 adherence of the toxin to the cell surface while the N-terminus
 mediates transport of the light chain from the endocytic vesicle
 to the cytosol. After translocation, the light chain (L)
 hydrolyzes the 197-Gln-1-Arg-198 bond in SNAP-25, thereby blocking
 neurotransmitter release. Inhibition of acetylcholine release
 results in flaccid paralysis, with frequent heart or respiratory
 failure.
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
 CC heavy chain (H).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- PHARMACOLOGICAL: Available under the name BOTOX (Allergan) for
 CC the treatment of strabismus and blepharospasm associated with
 CC dystonia and cervical dystonia. Also used for the treatment of
 CC hemifacial spasm and a number of other neurological disorders
 CC characterized by abnormal muscle contraction.
 CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of
 CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -1- DATABASE: NAME=BOTOX product information Web site;
 CC WWW="http://www.botox.com/index.jsp?hpproductinfo".
 CC -1- DATABASE: NAME=Protein Spotlight;
 CC NOTE=Issue 19 of February 2002;
 CC WWW="http://www.expasy.org/spotlight/articles/spl1019.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X52066; CAA36289.1; -
 DR EMBL; M30196; AAA23262.1; -
 DR EMBL; X92973; CAA63551.1; -
 DR EMBL; D67030; BAA11051.1; -
 DR EMBL; M27892; AAA23269.1; -
 DR PIR; A35294; BTCLAB.
 DR PIR; S09492; S09492.
 DR PDB; 3BTA; O1-OCT-99.
 DR MEROPS; M27.002; -
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
 KW Pharmaceutical; 3D-structure.
 FT INIT_MDT 0
 FT CHAIN 1 447
 FT CHAIN 448 1295
 FT METAL 222 222
 FT METAL 223 223
 FT METAL 226 226
 FT METAL 261 261
 FT DISULFID 429 453
 FT BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
 FT BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
 FT ZINC (CATALYTIC).
 FT ZINC (CATALYTIC).
 FT ZINC (CATALYTIC).
 FT INTERCHAIN.

```

FT DISULFID 1234 1279 POTENTIAL.
FT TRANSMEM 626 646 POTENTIAL.
FT TRANSMEM 655 675 V -> A.
FT VARIANT 26 26 E->A: DRASTIC DECREASE IN ENZYMAIC
FT MUTAGEN 261 261 ACTIVITY
FT MUTAGEN 265 265 F->A: DECREASE IN ENZYMAIC ACTIVITY.
FT MUTAGEN 365 365 Y->A: DECREASE IN ENZYMAIC ACTIVITY.
FT CONFLICT 1 1 P -> Q (IN REF. 1).
FT CONFLICT 479 479 E -> P (IN REF. 9).
FT CONFLICT 875 875 T -> L (IN REF. 8).
FT CONFLICT 891 891 S -> K (IN REF. 8).
SQ SEQUENCE 1295 AA; 149322 MW; 858342F75486579 CRC64;

Query Match 26.6%; Score 653.5; DB 1; Length 1295;
Best Local Similarity 33.1%; Pred. No. 3.2e-33;
Matches 163; Conservative 88; Mismatches 159; Indels 83; Gaps 18;

QY 4 STPIPEYSKMLDGVNEDIDVILK-----KSTLINDINDIISDSCGNSSVITY 57
D 845 STDIPOLSK---YVNOQLLSTFTEYIKNTWTSTILNYESNHLIDLSRVASKIIG 900
QY 58 PDAOLVPGINGKAHLVNNESSEVIYHKAMIEYNDMFNFTVSFWLRVPR-----VSASH 113
D 901 SKVNPDP-IDKNQIQLEFLLESSKIEVLKNAIYVNSMYENSTSEFWIRIPKFNISILN- 958
QY 114 LEQYGTNEYSTIISMKKHSISGSGWSYSLKGNMLIWTLKDSAGEVQITFR-----DLP 168
D 959 -----NEYTIINCENN-----SGWKVSLNYGELIWTLOTOELIKORVPRKYSOMINIS 1007
QY 169 DKFNAYLANKVAVITITDRIKSNALYINGVLMGSAETTGAIREDNNITLKDRCNNN 228
D 1008 D-----YL-NWMEVITITNKRNNKSYIYNGRLIDOKPISNIGNHNINMFKIDGCDT 1062
QY 229 NOYVIDKFRIFCKALNPKIEKLYSTLSTFLDEWGNPLRYDTEYLLIPVASSKDV 288
D 1063 HRYIWKYFENLFDKELNKEIKLDIDNOSNGILKDFWGDYQYDKPYMLNLXYPNRYV 1122
QY 289 QLNK--ITDYMTLTPASPYTGNCKLNIYFR-RLYNGLKRTIKRYTPNNELDSFVKSDEIK 345
D 1123 DVNNNGIGGYMYL-KGPRGSVWTNTIYLNSSLYRGTKFLIKKYASGKN-DNIVRNNDRYV 1180
QY 346 LVYSYNNNEHLYGYPRKDNAFENNLDRILRVGYNAPGIPLYKKMEAVNKLRLDK----- 397
D 1181 INVYVKNNEY-----RLATNASQAGYEKLTSLALEIPDVGNLSQVAVM 1222
QY 398 -----TYSYQALKLYDDKNASLGLVGTHTNGQIGNDPNRDILIASNMYFNHLK--DKIL 447
D 1223 KSKNDGJITNCKKRNLDONNGNDIGIFPH--QFNNIK--LVASNMVNRQIERSSRTL 1277
QY 448 GCDWYFVPTDEGM 460
D 1278 GCSWEFIPVDGM 1290

RESULT 3
BXG_CLOBO STANDARD; PRT: 1296 AA.
AC Q60393;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type G precursor (Ec 3.4.24.69) (BONT/G)
DE (Bontoxilysin G).
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=113 / 30;
RX MEDLINE=94092745; PubMed=8268233;

```

```

RA Campbell K., Collins M.D., East A.R.;
RT "Nucleotide sequence of the gene coding for Clostridium botulinum
RT (Clostridium argentineense) type G neurotoxin: genealogical comparison
RT with other clostridial neurotoxins.";
RL Biochim. Biophys. Acta 1216:487-491(1993).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE.
CC -1- CATALYTIC ACTIVITY: limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAPs or syntaxin. NO
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted (by similarity).
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X74162; CA52275.1; .
DR HESP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_mtpetidase.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTTOXILYSIN.
DR PRODOM: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR Neurotoxin; Hydrolase; Metalloprotease; Zinc.
FT INTT_MET 0
FT CHAIN 1 441
FT CHAIN 442 1296
FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 230 230 BY SIMILARITY.
FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 435 449 INTERCHAIN (PROBABLE).
SQ SEQUENCE 1296 AA; DC8E47E15F665C31 CRC64;

Query Match 26.2%; Score 644.5; DB 1; Length 1296;
Best Local Similarity 33.2%; Pred. No. 1.2e-32;
Matches 161; Conservative 82; Mismatches 183; Indels 59; Gaps 17;

QY 7 IFPS---YSKN-----LDCWVNDNEDIDVILKSTLINDINDIISDSCGNSSVITY 57
D 840 IPFDLSLYTKRDFILIQVFNYSN-----ISSNKLISLSTRGGLIDSSGGAIMNG 892
QY 58 PDAOLVPGINGKAHLVNNESSEVIYHKAMIEYNDMFNFTVSFWLRVPRKYSASHLEYQ 117
D 893 SDV-IFNDIGNQOFKLNNSNENSIYTAHQSKFYVYDSMFNFSINFWVRPKYNNNDIQY 951
QY 118 GNEYSIISMKKHSISGSGWSYSLKGNMLIWTLKDSAGEVQITFR-DLPKFNAYLA 176
D 952 LQNEYTIISCIKN-----DSGKWSIKNGRIITWLLIDVNAKRSIFEYSIKNIDSYL- 1005
QY 177 NKWVFTITINDRNSANLNTNGVLMGSAETTGAIREDNNITLKDRCNNNOYVSIK 236
D 1006 NKWFTITINDRNGANNIYIYNSLAKSEKILNDRINSSNDIDFKILNCIDTTRFWIKD 1065
QY 237 FRIFCKALNPKIEKLYSTLSTFLDEWGNPLRYDTEYLLIPVASSKDVOLKNTDY 296
D 1066 FNFIFGRELNATEVSLYQSSNTITLKDFWGNPLRYDYOYLF--NQGQNIYIKYFESKA 1123

```


ID BXR.CIOBO STANDARD: PRT: 1290 AA.
 AC P10844; P10843; (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BoNT/B)
 DE (Bontolysin B).
 CN BoNT.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 CX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92384550; PubMed=1514783;
 RA Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
 RA Minton N.P.;
 RA "Molecular cloning of the Clostridium botulinum structural gene
 RA encoding the type B neurotoxin and determination of its entire
 RA nucleotide sequence.";
 RT Appl. Environ. Microbiol. 58:2345-2354(1992).
 RL [2]
 RN SEQUENCE OF 35-245 FROM N.A.
 RP STRAIN=NCCTC 7273;
 RC Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 633-993 FROM N.A.
 RC STRAIN=NCCTC 7273;
 RX MEDLINE=94013372; PubMed=8408542;
 RA Campbell K., East A.R., Collins M.D.;
 RT "Gene probes for identification of the botulin neurotoxin gene and
 RT specific identification of neurotoxin types B, E, and F.";
 RL J. Clin. Microbiol. 31:2255-2262(1993).
 RN [4]
 RP SEQUENCE OF 1-44 AND 441-466.
 RC STRAIN=657;
 RX MEDLINE=89000987; PubMed=3139097;
 RA Dasgupta B.R., Datta A.;
 RT "Botulinum neurotoxin type B (strain 657): partial sequence and
 RT similarity with tetanus toxin.";
 RL Biochimie 70:811-817(1988).
 RN [5]
 RP SEQUENCE OF 1-16 AND 441-458.
 RC STRAIN=OKRA;
 RX MEDLINE=85197963; PubMed=3888113;
 RA Schmidt J.J., Satyamootthy V., Dasgupta B.R.;
 RT "Partial amino acid sequences of botulinum neurotoxins types B and
 RT E.";
 RL Arch. Biochem. Biophys. 238:544-548(1985).
 RN [6]
 RP IDENTIFICATION AS ZINC-PROTEASE.
 RX MEDLINE=93054654; PubMed=1428690;
 RA Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;
 RT "Botulinum neurotoxins are zinc proteins.";
 RL J. Biol. Chem. 267:23479-23483(1992).
 RN [7]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=93063293; PubMed=1331807;
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
 RA Dasgupta B.R., Montecucco C.;
 RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
 RT by proteolytic cleavage of synaptobrevin.";
 RL Nature 359:832-835(1992).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-1-PHE-77 BOND OF
 CC SYNAPTOSOMAL-2.
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO

CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 DR EMBL: M81186; AAA2321.1; -;
 DR EMBL: Z11934; CAA77991.1; -;
 DR EMBL: X70817; CAA50148.1; -;
 DR PIR: S07128; S07128;
 DR PIR: S07155; S07155;
 DR PIR: S08562; S08562;
 DR PIR: S08573; S08573;
 DR PIR: S08574; S08574;
 DR PIR: A48940; A48940;
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002; -;
 DR InterPro: IPR000395; Bontolysin.
 DR InterPro: IPR000130; Zn_Mpeptidse.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOLYSIN.
 DR PRODOM: PD001963; Bontolysin.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 KW INT_MET 0
 FT CHAIN 1 440
 FT CHAIN 1 1290
 FT METAL 229 229
 FT ACT_SITE 230 230
 FT METAL 233 233
 FT DISULFID 436 445
 FT CONFLICT 29 29
 FT CONFLICT 217 217
 FT CONFLICT 224 224
 FT CONFLICT 463 463
 SQ SEQUENCE 1290 AA; 150670 MW; D21746E2C024DF43 CRC64;
 Query Match 24.0%; Score 590; DB 1; Length 1290;
 Best local Similarity 31.7%; Pred. No. 2; 9e-29;
 Matches 140; Conservative 89; Mismatches 184; Indels 28; Gaps 12;
 QY 34 LNLINNDIISDGFNSVITYPPDAOLVPGINGK-AIHLVNNSESEVTVKAMDIEYN 92
 DB 864 LNLKRYKNDLIDSLGKAYEYDDEL-----NDKNQFLTSANSKINVTQNTIFN 919
 QY 93 DMENFTVSFWLVPKVSASHEOYSTNEYSISSMKHSLSGSGSVSLKGNLITWL 152
 DB 920 SVFLDSVSFWIRIPKXKNDGIONITHNEYTIINCKMNN-----SGWKISIRGRRIITWL 974
 QY 153 KKSAGVROITFR-DLPDFENATLANKWFEITTRNLSANLYINVLNGMSARITGGA 211
 DB 975 IDINCKTSVFEEYNIREDISERY-NKWFVTTINN-LNNAKITINCKLESNDIKDIRE 1032
 QY 212 IREDNNITLKCRCNNNOYSDIKFRIFKALNPKEIEKLYSYSLITELDFWGNPLR 271
 DB 1033 VYANGELIIRKLODDIDRTQFIMWKYSIFTELSSQSIIEERYKIQSYEVLKDFWGNPLM 1092
 QY 272 YTEYEYLLIPVASSKVOALKNITDYMVLITAPSYTNKLNIIYRRLYNGLKFLIKRTPN 331
 DB 1093 YNKEYYMFAGNKNNSYIKLKDKSPVEILTRSKYNQNSKYINFDLYIGEFILIRRSNS 1152


```

Db 822 IPRKLSYTD-----DKILISYFNKFFRIKSSSVLMNRKNDKYVDTSYDSSMIN 872
OY 56 TYPDAQVPGINGKAIHLVNNSSSEVYIHKAMDIEYNMFNNFTVSEFWLRVKSASHLE 115
Db 873 INGVADKYYP-TNRKNQFGIYNDKLSSEVNISSNDYIILYDNKNKFSISFWRIPIYNTKIVN 931
OY 116 OYGTVEYSIISMKHSHSISGMSVSLKGNLITWLKDSAGEVBOITFR-DLPDKFNAY 174
Db 932 V--NNEYITINMRNN-----SGWKVSLNHNELIWTPEENRGINOKLAFVNGNANGISDY 985
OY 175 LANKWVFITNDRLSSANLYNGVLMGSAETTGAGIREDNNITLKDRCNNNNQYVS 234
Db 966 I-NKMFVITITNDRLGDSKLYINGNLIDOKSLINLGNHVSNDLIFKLYNC-SYTRYIGI 1043
OY 235 DKFRICFALNKEIEKLEKLTSTFLRDFGPNPLRYDTEYLLIPVASSS-----KD- 287
Db 1044 RYFNIFDKELDETEIOTLTSNEPNINILKDFGNNILLYKEVYLLNLVLPKPNFTDRKDS 1103
OY 288 -VOLKNITDYMYLTNAPSYNGKLNITYRRLYNGLEIKRYTPNNEIDSEYKSGDFITL 346
Db 1104 TISINNIRSTILLAN-----RLYSGIKVKIQRVNNSSTNDNLVKKND--OV 1147
OY 347 YVSY-----NNNEHYGYPKDGNFNNLDRILRVGNAPGIFLYKKMEA 390
Db 1148 YINFAVSKTHLPYADYATNTKEKTIKISSGGRNOV-----VYANSVG 1193
OY 391 VKLRDLKTYSVOLKLYDKNASLGLVGTNGIGDNDPNDILASWYFNHLEKD--KILG 448
Db 1194 -----NCWTFNFKNNNGNINIGLGF-----KADTVVASTWYVYHMDRHTNSNG 1235
OY 449 CMWYFVPTDEGM 460
Db 1236 CFWNFISEHGM 1247

RESULT 7
BSE_CLOBU STANDARD; PRT: 1250 AA.
ID BSE_CLOBU STANDARD; PRT: 1250 AA.
AC P30995;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)
DE (Bontoxilysin E).
OS Clostridium butyricum
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1492;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43181, and ATCC 43755;
RX MEDLINE=92181428; PubMed=1543481;
RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT "Sequences of the botulinum neurotoxin E derived from Clostridium
botulinum type E (strain Beluga) and Clostridium butyricum (strains
ATCC 43181 and ATCC 43755).";
RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
RN [2]
RP SEQUENCE OF 1-251 FROM N.A.
RC STRAIN=BL6340;
RX MEDLINE=91237316; PubMed=2033376;
RA Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
RA Yokosawa N., Yashiki T., Oguma K.;
RT "Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
type E toxin gene from Clostridium butyricum strain BL6340.";
RL J. Gen. Microbiol. 137:519-525(1991).
RN [3]
RP SEQUENCE OF 1-48.
RC STRAIN=5262;
RA Gimenez J., Foley J., Dasgupta B.R.;
RT "Neurotoxin type E from Clostridium botulinum and C. butyricum:
partial sequence and comparison.";
RL FASEB J. 2:A1750-A1750(1988).

```

```

CC -I- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE.
CC -I- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -I- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -I- SUBCELLULAR LOCATION: secreted.
CC -I- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X62088; CAA43998.1; -
CC EMBL: X53180; CAA37321.1; -
CC PIR: JH0256; JH0256.
CC PIR: S16145; S16145.
CC HSP: P10845; 3BTA.
CC MEROPS: M27.002; -.
CC InterPro: IPR000395; Bontoxilysin.
CC InterPro: IPR000130; Zn_MTPeptidse.
CC Pfam: PF01742; Peptidase_M27; 1.
CC PRINTS: PR00760; BONTOTOXILYSIN.
CC ProDom: PD001963; Bontoxilysin; 1.
CC PROSITE: PS00142; ZINC_PROTEASE; 1.
CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
CC INIT_MET 0 421
CC CHAIN 1 421
CC CHAIN 422 1250
CC METAL 211 211
CC ACT_SITE 212 212
CC METAL 215 215
CC DISULFID 411 425
CC CONFLICT 229 229
CC SEQUENCE 1250 AA; 143265 MW; 8171B5B2C312857 CRC64;

Query Match 22.8%; Score 560.5; DB 1; Length 1250;
Best Local Similarity 29.3%; Pred. No. 1.9e-27;
Matches 147; Conservative 83; Mismatches 164; Indels 107; Gaps 17;

OY 1 MWFST---PIPFYSKNDLCWVNEEDIDVL-----KSTILNDINNDITSD 46
Db 813 MVIDLNNISIFKLSSTYD-----DKILISYFNKFFRIKSSSVLMNRKNDKYVD 863
OY 47 ISGENSVYITTPDAQVPGINGKAIHLVNNSSSEVYIHKAMDIEYNMFNNFTVSEFWLRV 106
Db 864 TSGYDSNININGDYKYP-TNRKNQFGIYNDKLSSEVNISSNDYIILYDNKNKFSISFWRI 922
OY 107 PKVASHEOYGTNEYSIISMKHSHSISGMSVSLKGNLITWLKDSAGEVBOITFR- 165
Db 923 PRYNDKRYNV--NNEYITINMRNN-----SGWKVSLNHNELIWTPEENRGINOKLAFV 976
OY 166 DLPDKFNAYLANKWVFITNDRLSSANLYNGVLMGSAETTGAGIREDNNITLKDRC 225
Db 977 GNANGISDYI-NKMFVITITNDRLGDSKLYINGNLIDOKSLINLGNHVSNDLIFKLYNC 1035
OY 226 NNNNGYVSDKFRICFALNPKIEKLTSTFLRDFGPNPLRYDTEYLLIPVASS-- 283
Db 1036 -SYTRYIGIRYNTDKELDETEIOTLTSNEPNINILKDFGNNILLYKEVYLLNLVLPKPN 1094

```



```

OY 284 -----SSKDVOLKNITDYMVLINAPSTYNGKLNLYRRRLNGLKFTIKRYPNNEIDSF 337
DB 1095 NFINRRDSTLSINNIRKSTILLAN-----RLXSGIKVKIQRVNWSSTNNL 1140
OY 338 VKSGDFELKLVSY-----NNNEHIVGPKRGNAFNNIDRLIRVGNAPG 381
DB 1141 VRKND--QVYINFAVSKTHLLPYADATTTAKETIKISSGNNRNOV-----VYVNSVSG 1193
OY 382 IPLYKKEAVKRLDKTYKSVOLKLYDDKNASLGVTGHNQIGNDPNHDIILASWYFNH 441
DB 1194 -----NCTMNFKNNNNGNIGLLGF-----KADIVYASWYTYTH 1226
OY 442 LKDKI--LGCDWYFVPTDEGM 460
DB 1227 MRDNTNSNGEFMNFISEHGM 1247

RESULT 8
BXF_CLOBO STANDARD; PRT; 1274 AA.
ID BXF_CLOBO P30996;
AC 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BoNT/F)
DE (Bontoxilysin F).
GN BOPF.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OC NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23387;
RA MEDLINE=93012902; PubMed=1398040;
RA East A.K., Richardson P.T., Allaway D., Collins M.D.,
RA Roberts T.A., Thompson D.E.;
RT "Sequence of the gene encoding type F neurotoxin of Clostridium
RT botulinum.";
RL FEWS Microbiol. Lett. 75:225-230(1992).
RN [2]
RP SEQUENCE OF 1-64 FROM N.A.
RC STRAIN=Hobbs FT10;
RA MEDLINE=94297488; PubMed=7764998;
RA East A.K., Collins M.D.;
RT "Conserved structure of genes encoding components of botulinum
RT neurotoxin complex M and the sequence of the gene coding for the
RT nontoxic component in nonproteolytic Clostridium botulinum type F.";
RN [3]
RP CURR. Microbiol. 29:69-77(1994).
RN [3]
RP SEQUENCE OF 634-1002 FROM N.A.
RA MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RN [4]
RP J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP IDENTIFICATION OF SUBSTRATE.
RA MEDLINE=94230352; PubMed=8175689;
RA Yamasaki S., Baumeister A., Binz T., Blas J., Link E., Cornille F.,
RA Roques B., Eykse E.M., Suedhof T.C., Jahn R., Niemann H.;
RT "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT F botulin neurotoxins and tetanus toxin.";
RN [5]
RP J. Biol. Chem. 269:12764-12772(1994).
RN [5]
RP -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
RN RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
RN AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
RN WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
RN INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
RN ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-1-LYS-59
RN BOND OF SYNAPTOBREVIN-1 AND -2.
RN -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
RN neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO

```

```

CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.sib.ch).
CC -----
DR EMBL; M92906; AAA23263.1; -
DR EMBL; S73676; AAC60475.1; -
DR EMBL; X70820; CA50151.1; -
DR EMBL; X70816; CA50147.1; -
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_MPeptidase.
DR Pfam; PF01742; Peptidase_M27_1.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR PRODOM; P0001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR Neurotoxin; Transmembrane; Hydrolyase; Metalloprotease; Zinc.
FT CHAIN 1 436
FT CHAIN 2 1274
FT METAL 227 227
FT ACT_SITE 228 228
FT METAL 231 231
FT DISULFID 429 445
SQ SEQUENCE 1274 AA; 146709 MW; 5B99756A7438B921 CRC64;

Query Match 21.1%; Score 517; DB 1; Length 1274;
Best local similarity 29.1%; Pred. No. 9,9e-25;
Matches 131; Conservative 99; Mismatches 160; Indels 60; Gaps 17;

OY 29 LKSGTILNLDINNDIISDPSGFSVITYPPDAQLVGINGKA-----HLVNNSS 79
DB 865 IKDSIILDMRENNKFIDISGYSNI-----SINGNVIYISTNNRGITYNSRLS 914
OY 80 EYIYHKAMDIENMFNFYSEMLRVRKVSASLEOYGTN-EYLSISMKRHSLSIGSG 138
DB 915 EVNIAQNNDIYNSRYONFISFWRIK---HYKPMNHREYTLINCNGN---SG 966
OY 139 WSVSL---KGNLIWTLKDSAGEVQIIFR-DLPKFNAYILANKVETITDRSSANL 194
DB 967 WKISLRVROCEIITWTDTSCKENLIFRYELNRISYTI-NKMEIVITTNRLGNSRI 1025
OY 195 YINGVLMSAEITGALIREDNNTTLKDRCNMNNQYVSDKFRFCALNKEIEKLYT 254
DB 1026 YINGNLIVKSSISNIGDIHVDNLIFKIVGDDER-VYGIIRFKFENELDTETETITS 1084
OY 255 SYLSITFLRQWGNPLRYDTYYLLPVAASSKDVOLKNTIDYMYTLNAPSTYNGKLNYY 314
DB 1085 NEPDPSILKNWGVLLYNNKYVLENLKRDKYITLN--SGILNINOORCVEGSVFLNY 1142
OY 315 RLYNLGKFIKRYTP--NNEIDSFVKSGDFTIKLYSVSNNEHIVGPKRGNAFNNIDRI 372
DB 1143 -KLYEGEVALIRKNGPDISDITDMFVRKNDL--AYI--NVYDREYRLYADTKSEKEKI 1197
OY 373 LRYGYNAPGIPLYKKEAVKRLDKTYKSVOLKLYDDKNASLGVTGHNQIGNDPNHDI 432
DB 1198 IRYS-----NLNDSIQIILYMDISGNCTNPNNGNSNIGLLGFHSNN-----L 1242
OY 433 IASNWFENHLKDKIL--GCDWYFVPTDEGM 460

```

Db 1243 VASSWYNNIRNTSSNCCFWSSISKENGW 1272

RESULT 9

CC BXCL_CLOBO STANDARD: PRT: 1290 AA.

AC P18640.

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

DE Botulinum neurotoxin type C1 precursor (EC 3.4.24.69) (BoNT/C1)

DE (Bontoxilysin C1).

OS Clostridium botulinum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI_TaxID=1491;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90370487; PubMed=2204031;

RA Hausner D., Eklund M.W., Kurazono H., Binz T., Niemann H., Gili D.M.,

RT Boquet P., Popoff M.R.;

RL "Nucleotide sequence of Clostridium botulinum C1 neurotoxin.";

RL Nucleic Acids Res. 18:4924-4924(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX STRAIN=Type C, Stockholm / C-ST;

RX MEDLINE=91024998; PubMed=2222445;

RA Kimura K., Fujii N., Tsuzuki K., Murakami T., Indoh T.,

RA Yokosawa N., Takeshi K., Syuto B., Oguma K.;

RT "The complete nucleotide sequence of the gene coding for botulinum

RT type C1 toxin in the C-ST phage genome.";

RL Biochem. Biophys. Res. Commun. 171:1304-1311(1990).

RN [3]

RP SEQUENCE OF 2-25.

RX STRAIN=Type C Stockholm / C-ST;

RX MEDLINE=88153072; PubMed=2450068;

RA Tsuzuki K., Yokosawa N., Syuto B., Ohishi I., Fujii N., Kimura K.,

RA Oguma K.;

RT "Establishment of a monoclonal antibody recognizing an antigenic site

RT common to Clostridium botulinum type B, C1, D, and E toxins and

RT tetanus toxin.";

RL Infect. Immun. 56:898-902(1988).

RN [4]

RP IDENTIFICATION OF SUBSTRATE.

RX MEDLINE=94038966; PubMed=7901002;

RA Blas J., Chapman E.R., Yamasaki S., Binz T., Niemann H., Jahn R.;

RT "Botulinum neurotoxin C1 blocks neurotransmitter release by means of

RT cleaving HSC-1/syntaxin.";

RL EMBO J. 12:4821-4828(1993).

CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER

CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED

CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD

CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT

CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC

CC ENDOPEPTIDASE THAT CLEAVES SYNAPSIN.

CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the

CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No

CC detected action on small molecule substrates.

CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A

CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,

CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL

CC FORMATION AND TOXIN BINDING, RESPECTIVELY.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF

CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.

CC -1- MISCELLANEOUS: BOTULINUM TYPE C1 NEUROTOXIN IS SYNTHESIZED BY C

CC STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE

CC BACTERIOPHAGE.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: X66433; CAA47060.1; -

DR EMBL: X72793; CAA51313.1; -

DR EMBL: X53751; CAA37780.1; -

DR EMBL: D90210; BAA14235.1; -

DR EMBL: X62389; CAA44263.1; -

DR PIR: S11291; S11291.

DR PIR: A35396; A35396.

DR PIR: A43503; A43503.

DR HSSP: P10845; 3BTA.

DR MEROPS: M27.002; -

DR InterPro: IPR000395; Bontoxilysin.

DR InterPro: IPR000130; Zn_Mpeptidase.

DR Pfam: PF01742; Peptidase_M27; 1.

DR PRINTS: PR00760; BONTOTOXILYSIN.

DR ProDom: PD001963; Bontoxilysin; 1.

DR PROSITE: PS00142; ZINC_PROTEASE; 1.

KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.

FT INIT_MET 0 0

FT CHAIN 1 448 BOTULINUM NEUROTOXIN C1, LIGHT-CHAIN.

FT METAL 449 1290 BOTULINUM NEUROTOXIN C1, HEAVY-CHAIN.

FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).

FT ACT_SITE 229 229 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).

FT DISULFID 436 452 INTERCHAIN (PROBABLE).

FT CONFLICT 84 84 P->T (IN REF. 2).

SQ SEQUENCE 1290 AA; 148734 MW; 71FB379F97129E8 CRC64;

Query Match 17.5%; Score 430; DB 1; Length 1290;

Best Local Similarity 26.4%; Pred. No. 2.6e-19;

Matches 135; Conservative 90; Mismatches 172; Indels 114; Gaps 21;

QY 3 FSTPIPFYSKLNDCVNDIEDIVT-----LKKSTILNLDINNDISISGFNSSVYT 56

Db 839 FONTIPF-----NIFYSTNNSLKDIINEYFNNDISKILSLQRRKNTLVDTSSYNAEVS 894

QY 57 YPDQOLVP-----GINKKATHLVNNESEYVHKRAMDIENDMFNNTVGFMLRVPK 108

Db 895 EGVQOLNPIPFDPFKLGSSG-----DGKVIYQNEINVIYNSMESISIPWRINK 947

QY 109 VSASHLEOYGTNEYSLISSMKKHSLSIGGSVSLGNNLIWTLKDSAGEVROITFR-DL 167

Db 948 -WVSNLP-----GYTIIDSVKNN-----SGMSIGIISNFIPLFKQNEDESGSINFSDYDI 996

QY 168 PDKFNAYLANKWVFITITNDRSSANLYINGVLMGSAELITGLAIRDNNTITLKLDR--- 224

Db 997 SNNAPGY--NKKWFYVITNNMGMMKLYINGKLIDTLKVELTGINSKTIITFEINKIPD 1054

QY 225 -----CNNNQVYSIDKFRIFCKALNKEIEKLYTSLSTFLDFEFGNPLRYDTEYLI 279

Db 1055 TGLITSSDNIINMIRFIFYFAKELDDKDLINLFNSLOQYINNVAVDYGNDLRKVEYMW 1114

QY 280 PVASSKDVOLKNTDYMVLYLNAFSTNGKLNITYRRLYN-----GLKFIITKRTPPNEI 334

Db 1115 NI-----DYL-----NRYMANSROIYFNTRNNNDNEGKIIITKIRLGWTN- 1157

QY 335 DSYVKSDFIKLYSYNNNEHIVCPDGNAFNNLDILIRGVYAPGPIYKKKEAAVK-- 392

Db 1158 DTRVRGDIILYFDMTINKAV-----NLEMKNETMYADNISTEDIIYAIGLRQIKDI 1209

QY 393 -----LRDLKTSYVQ--KLYDDKNAS-GLVGTGNGQIGNDPNRDIILASWY- 438

Db 1210 NDNITIPQIPANNIYYASQIFKFNNGENSGICSGTYFRFGD-----WYR 1259

QY 439 FNHLKDKI-----LGCQWYFVPTDE 458

Db 1260 HNYLVPVYKQGNVYASLLESTSTHWGFVPSV 1290

RESULT 10
 BND_CLOBO STANDARD; PRT; 1276 AA.
 ID BND_CLOBO STANDARD; PRT; 1276 AA.
 AC P19321;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DDE Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BONT/D)
 DE (Bontolysin D).
 GN BONT.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BVD-3;
 RA MEDLINE=91016853; PubMed=2216736;
 RA Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,
 RA Kozaki S., Krieglstein K., Henschen A., Gill D.M., Niemann H.,
 RA "Nucleotide sequence of the gene encoding Clostridium botulinum
 RL neurotoxin type D.";
 RL Nucleic Acids Res. 18:5556-5556(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CB16;
 RA MEDLINE=93042276; PubMed=1420572;
 RA Sunagawa H., Ohnaya T., Watanabe T., Inoue K.;
 RA "The complete amino acid sequence of the Clostridium botulinum type D
 RL neurotoxin, deduced by nucleotide sequence analysis of the encoding
 RL phage d-16 phi genome.";
 RL J. Vet. Med. Sci. 54:905-913(1992).
 RN [3]
 RP PARTIAL SEQUENCE.
 RC STRAIN-D-SA, and D-1873;
 RX MEDLINE=893939741; PubMed=2668193;
 RA Morishiki K., Syuto B., Kubo S., Oguma K.;
 RA "Molecular diversity of neurotoxins from Clostridium botulinum type D
 RL strains.";
 RL Infect. Immun. 57:2886-2891(1989).
 RN [4]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94230352; PubMed=8175689;
 RA Yamasaki S., Baumeister A., Binz T., Blas J., Link E., Cornille F.,
 RA Rogues B., Fyke E.M., Suedhof T.C., Tann R., Niemann H.;
 RA "Cleavage of members of the synaptobrevin/VAMP family by types D and
 RL F botulinum neurotoxins and tetanus toxin.";
 RL J. Biol. Chem. 269:12764-12772(1994).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES THE 60-LYS-1-LEU-61 BOND OF
 CC SYNAPTOSOMAL-1 AND -2.
 CC CATALYTIC ACTIVITY: limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO
 CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC MISCELLANEOUS: BOTULINUM TYPE D NEUROTOXIN IS SYNTHESIZED BY D
 CC STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
 CC BACTERIOPHAGE.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X54254; CAA38175.1; -
 CC EMBL: S49407; AAB24244.1; -
 CC PIR: S11455; S11455.
 CC HSSP: P10845; 3BTA.
 CC MEROPS: M27.002; -.
 CC InterPro: IPR000395; Bontolysin.
 CC InterPro: IPR000130; Zn_Mpeptidase.
 CC Pfam: PF01742; Peptidase_M27; 1.
 CC PRINTS: PR00760; BONTOLYSIN.
 CC PRODOM: PD001963; Bontolysin.
 CC PROSITE: PS00142; ZINC_PROTEASE; 1.
 CC KEGG: Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 CC CHAIN 1 442
 CC METAL 443 1276
 CC ACT_SITE 229 229
 CC ACT_SITE 230 230
 CC METAL 233 233
 CC DISULFID 437 450
 CC VARIANT 15 16
 CC VARIANT 17 18
 CC VARIANT 452 452
 CC VARIANT 457 457
 CC VARIANT 457 457
 CC VARIANT 457 457
 CC VARIANT 462 462
 CC VARIANT 489 489
 CC VARIANT 644 644
 CC VARIANT 1122 1122
 CC SEQUENCE 1276 AA; 146871 MW; C1EC50F46C8233E2 CRC64;
 Query Match 16.88; Score 411.5; DB 1; Length 1276;
 Best Local Similarity 24.9%; Pred. No. 3,7e-18;
 Matches 127; Conservative 96; Mismatches 164; Indels 123; Gaps 19;
 QY 3 FSPPIPSYKSLNDCWDNEDIDY-----LKKSTLNDINDNDISDGSFSSVTT 56
 DB 836 FENIMPF-----NISTYNNSLKIDTINEYNSINDSKILSLQNKNAVLVDYSGYNAEYV 891
 QY 57 YPDALVPGINGKAIRH-----VNNSESEVIAHKAMDIEYNMFNFVSPFLVRPK-USA 111
 DB 892 GDNVOL-----NTIYNDPFLSSGDKIYLNANNILYALYENSVSFWTKISKDLTN 945
 QY 112 SHLEQGTNEYSTIISKKSSLSIGSGWSVLKGNLITWLKDSAGEVQITFRDLPDKF 171
 DB 946 SH-----NEYTIINSLEON-----SQMKLCIRNGNIEWIIQDVARKKYSILFDYSESLS 994
 QY 172 NAYLANKWVEITTTNDRLSSANLYINGVLMGSAEITGATRENNITLKLDRCNNNOY 231
 DB 995 HCTGYTKKWEFTITNNIMGIMKLYINELKOSQKIEDLEKLVKTYVFGIDENIDENOM 1054
 QY 232 VSDIKRIFPKALNPKETIEKLYTSYSITFLRDPGNDPLRYDEYLLIPVASSKQVOLK 291
 DB 1055 LWRIDNIEFSKESLNEDINIVYEQILRNVIKDYWGMDPKEDTEYIIN----- 1103
 QY 292 NLIDYVLTNPASVTNGKNIYY---RLYNGLKFTIKRTIPNNEIDSVFVSGDFIKLYV 348
 DB 1104 --DNYIDRIYAP-SNVLVLYOYPPDRSKLYTGNPTIKSVSDKNPY-SRLINGNIIILHM 1159
 QY 349 SYNNNEHIVGPKDGNAFNNLDRILRGYNAPGIPLYKKMEAYVLRDKT-YSVO----- 402
 DB 1160 LKNSRKMYL-----IROTDTIYAIOGGEGCS 1184
 QY 403 -----LKYLD-KNASLGLVGTN-----GQIGDNDNRDILLASN---WYFN--- 440
 DB 1185 QNCYVALKLOSNNGNGIGIFSIKNTIVSKNKYCSQFSFRENMTLADYIKPWFSPFKN 1244
 QY 441 -----HLKDKITGCD--WYFVPTDEGW 460
 DB 1245 ATPVAVNTYETKLLSTSSFWKTSRDPGW 1274


```

FT CARBOHYD 370 370 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 480 480 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 556 556 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 595 595 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 610 610 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 697 697 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 850 850 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 976 976 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1070 1070 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1087 1087 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1100 1100 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1112 1112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1117 1117 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1134 1134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1147 1147 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1225 AA; 134728 MW; 1D7650D89A48C9D CRC64;

```

Query Match 5.5%; Score 134.5; DB 1; Length 1225;
 Best Local Similarity 18.1%; Pred. No. 0.61;
 Matches 101; Conservative 85; Mismatches 180; Indels 193; Gaps 27;

```

QY 6 PTP-PSYKNDLCWDNED-----IDVIL-----KSTJLNDIND 42
DB 83 PPSFEAASSTCEGADPCNGAVLNTVDIRENFTTNVSGKATVSLTWTG 142
QY 43 ITSDISGFN-----SSVITYPDQALVP-GI-NGKAIHLVNNSESEVIVAKAMDIENDF 95
DB 143 VTEICSDYDSDSSFSYGE---IPGVITNPRCYV-----LY 180
QY 96 NNFVSFWLRVP---KVSASHLEOYGTMEYSITSSMKKHSISGWSYSLKGNLIMT 151
DB 181 NGTALVLTGLTPSVKEIATSKWGHFYINGYNFSTFPIDCIS---FNLITGDSDFWT 236
QY 152 LKSAEVRQITRDLRDEKFNAYLANKWFEITTNDRSSANLYINGVLSAETIGLA 211
DB 237 I-----ATTSTYEALOYENTAITVT-YCHSYV----- 264
QY 212 IREDNNITTLKDRCONNOVSIIDKFRIFCKALNPKIEKLYSYSLTFL----- 262
DB 265 ---NNIKCSQULANLNGFYPSSEV--GSYKNSYV--LPSFLHITITVITIGLGM 316
QY 263 RDMWGNPLRDTREYLYLPAVSSSKDYOL-----FNITDYM- 297
DB 317 RSGYGPDIASFTSLNITLPMQDNNITDVCVRSDFSVYVHSTCSALMDNFKRNCIDVD 376
QY 298 ---YLTNASTYNGKLNITYRRLYN-----GLKEIIRKRYTPNNEIDSEVKS 342
DB 377 ATAVIKGTGCPFSFDKLNLYL--TFNKFCLSLSPVGANCKEFDVAARRTN--DOFVRS-- 430
QY 343 FIKLYSYNNNEHIVGPKGMAFNNDRLRLRG-----YANPGLPYKKMAVAVLRD 395
DB 431 ---LYVIYEEGDSIVGYPSONSGIHLDS--VLHDSCTDYNIGRTGIGILRQTNRLLSG 486
QY 396 LKTYSVQLKLYDOKNASTGLV-----GTHNGOIGNPNNDILASNM-- 437
DB 487 LYVTSLSGDLGFKNVSDEGIYVTPCDVSAQAALIDGALVGA--TSINSBLALHTMTI 545
QY 438 -----YFNHLKDKILG 448
DB 546 TPNFYYSIVYNTDKTRG 564

```

RESULT 15
 P100_HSV7J STANDARD; PRT; 755 AA.
 ID P100_HSV7J
 AC P52519;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Large structural phosphoprotein homolog (Pp100).
 GN U11.

```

OS Human herpesvirus (type 7 / strain J1) (HHV7).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=57278;
RN [1]
RP SEQUENCE FROM N.A.
RA Nicholas J.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
-i SIMILARITY: TO THE LARGE STRUCTURAL PHOSPHOPROTEINS OF HSV-6 AND
HCNV UL32.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: 043400; AAC54672.1;
KW Matrix protein: Phosphorylation.
SQ SEQUENCE 755 AA; 86580 MW; 4083744CC3F91DA CRC64;

```

Query Match 5.3%; Score 131; DB 1; Length 755;
 Best Local Similarity 22.3%; Pred. No. 0.54;
 Matches 110; Conservative 71; Mismatches 194; Indels 118; Gaps 27;

```

QY 44 ISD-----ISGNSSVITP--DAQVPGI-----NCKAIH--LVNNESEVIVHK 85
DB 11 ISDEAKCFLSREFENISSLEPYVDIRENFWLSOCYKTKGINSNNVTLVNNLIMLYFHQ 70
QY 86 AM---DIEYNDMFNFTVSEFWLRVPVSAHLEQ-YGTMEYSITSSMKKHSISGWSY 141
DB 71 TLCKKRPDIHEWQEL-----LKQKILKQYLRQPMITDYSLSFSNK-----VGFEY 119
QY 142 SLK-----GNLITWLKDSAGEYRQITFRDLP-----KENAY-LANK 178
DB 120 EFKVNAKDLKLGSPFRMGVTHAADVYVNLTEERAIEGNLOKAKNNMLSTFIYQIVDP 179
QY 179 W-----VEITTNDRSSANLYINGVLMGS---AEITGLAIREDDNITTLKDRCONNOY 231
DB 180 WENEGYVVTINKLYLIGNLLT--TLHGSMMNNEKALANTINEKKNILK--ATENKNE 235
QY 232 VSIDKFRIFCKALNPKIEKLYSYSLTFLRDFWGNPLKDTREYLYLPAVSSSKD---- 287
DB 236 VSIVSYQILSLPLTSHRY---TSFEKI-LTEDEP--DVITKSLHLALPVKSTWDHVKF 288
QY 288 ---VQ-LKNITDYMVLTNASTYNGKLNITYRRLYNGLKFIIRKRYTPNNEI--DSFVKS 340
DB 289 TPEDIOEFKVLTDLSKSLSNQFSSSKTKSHGSPFPEPIKTEQORSNNTLSKDLVGS 348
QY 341 GDFIKLYSYNNNEHIVGPKGMAFNNDRLRLRGVGNAPGIPLYKKMAVAVLRDLYTS 400
DB 349 ED--GLSSVYKQSMILDEPRNSTSINNSKMHRI-----LQTEIIDLDTQTMHR 396
QY 401 VOLKLY-----DKNKNA--SLGVGHNGOIGNDPNDILIASWVENHL--KDK 445
DB 397 PEDKVNQFNEIYAAPDGINOVITFLSKLDLHNS-----NKVIDIVSPKVVYQLEKKN 450
QY 446 ILGCDWYFVPTDE 458
DB 451 IDVHSTFLEPENE 463

```

Search completed: December 1, 2002, 11:22:06
 Job time : 35 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2002, 11:16:52 ; Search time 80 Seconds

(without alignments)
1192.498 Million cell updates/sec

Title: US-09-816-467-2

Sequence: 1 MVESTRIPFSYSKNLDCWVD.....DKILCCDWYVPDTEGWTND 463

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTRMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_rvirus:*
17: sp_bacteriaph:*
18: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2359	96.1	1310	2	093N27
2	2340	95.3	451	2	091A13
3	617	25.1	1291	2	008077
4	600	24.4	1291	2	093K0
5	599	24.4	1291	2	092A58
6	597	24.3	1291	2	093K71
7	571	23.2	1268	2	045851
8	570.5	23.2	1251	2	093K95
9	567	23.1	441	2	09X708
10	565	23.0	1255	2	09FAR6
11	553	22.5	1278	2	057236
12	532.5	21.7	1280	2	092A55
13	430	17.5	1291	2	093HT3
14	417.5	17.0	1280	2	091B57
15	417.5	17.0	1280	2	045849
16	411.5	16.8	1275	12	09QTC7

17	407	16.6	1285	2	091BR1
18	406	16.5	1285	2	045967
19	274.5	11.2	1197	2	033871
20	274.5	11.2	1197	2	071117
21	269	11.0	1197	2	092A59
22	262	10.7	1197	2	045888
23	253.5	10.3	1196	2	091BR2
24	252.5	10.3	1196	2	091BR2
25	246.5	10.0	1196	2	091BR5
26	246.5	10.0	1196	9	092X77
27	246.5	10.0	1196	9	045893
28	245.5	10.0	1196	2	069276
29	244.5	10.0	1196	2	045916
30	244.5	10.0	1196	2	093HT4
31	244.5	10.0	1196	2	038197
32	243	9.9	1162	9	092A56
33	241	9.8	361	2	045846
34	241	9.8	361	2	045848
35	239	9.7	1198	2	006018
36	228	9.3	1160	2	087710
37	225	9.2	1193	2	045914
38	225	9.2	1193	2	069276
39	224	9.1	1161	2	071107
40	222	9.0	1193	2	045880
41	220	9.0	1163	2	045850
42	217.5	8.9	367	2	045861
43	217.5	8.9	367	2	045862
44	217	8.8	1161	2	045891
45	217	8.8	1193	2	071108

ALIGNMENTS

RESULT 1

093N27 PRELIMINARY; PRT; 1310 AA.
AC Q93N27;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Tetanus toxin (Fragments).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_Taxid=1513;
RN [1]
RP SEQUENCE FROM N.A.
RA Shumin Z., Dianliang L.;
RT "Cloning and sequence analysis of tetanus toxin gene."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF389424; AAK72964.2; -
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR001064; Crystallin.
DR InterPro: IPR000130; Zn_M7peptidase.
DR Pfam: PF01742; Peptidase_M27; 1.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
FT PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 1310 AA; 150316 MW; 9EADDG91418E450 CRC64;

Query Match 96.1%; Score 2359; DB 2; Length 1310;
Best Local Similarity 98.0%; Pred. No. 8.5e-132;
Matches 447; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 VESTRIPFSYSKNLDCWVDNEDIDVILKSTLTNDINNDISDISGFSYITPDAQ 61
Db 855 VESTRIPFSYSKNLDCWVDNEDIDVILKSTLTNDINNDISDISGFSYITPDAQ 914
Qy 62 LVPGNGKAHIVNNESSFVTHKAMDITENMENNFTYSFWLRVRKVASHLQYGTNE 121
|||||

```

Db 915 LVPINGKAIHLVNESSEVIHVKAMDEYNMFNFVFWLVPVSAHLEOYPTNE 974
OY 122 YSIISSKKHSLTSGSGMSVSLKGNLITWLKDSAGEVROITFRDLDPKFNAYLANKWV 181
    |||||
Db 975 YSIISSKKHSLTSGSGMSVSLKGNLITWLKDSAGEVROITFRDLDPKFNAYLANKWV 1034
OY 182 ITTINDRLSSANLVINGVLMGSAEITGAIREDNNITLKLDRCNNNOVYSIDKFRIFC 241
    |||||
Db 1035 ITTINDRLSSANLVINGVLMGSAEITGAIREDNNITLKLDRCNNNOVYSIDKFRIFC 1094
OY 242 KALPKKEIKLYTSLSTLFRDVGWGNPLRYDTEYLLIPVASSKDVOLKNTTYMYLTN 301
    |||||
Db 1095 KALPKKEIKLYTSLSTLFRDVGWGNPLRYDTEYLLIPVASSKDVOLKNTTYMYLTN 1154
OY 302 APSTYNGKLNITRYRLYGLKFEIKRYTPNNEIDSFVKSQDEIKLYSVNNNEHIVGYPK 361
    |||||
Db 1155 APSTYNGKLNITRYRLYGLKFEIKRYTPNNEIDSFVKSQDEIKLYSVNNNEHIVGYPK 1214
OY 362 DGMNFNNIDRLRVGNAPGIPLYKKMEAVKLRLDKTYSVOLKLYDDKNASGLVGTGNG 421
    |||||
Db 1215 DGMNFNNIDRLRVGNAPGIPLYKKMEAVKLRLDKTYSVOLKLYDDKNASGLVGTGNG 1274
OY 422 QIGNDPNRDILIASNMVFNHLKDKLTLCGDWYFVPTD 457
    |||||
Db 1275 QIGNDPNRDILIASNMVFNHLKDKLTLCGDWYFVPTD 1310

```

RESULT 2

```

O9LA13 PRELIMINARY: PRT: 451 AA.

```

```

AC 09LA13:
AC 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Tetanus toxin (Fragment).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxId=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=20886;
RA He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;
RT "Fragment C of Tetanus Toxin.";
RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF154828; AAF73267.1; -.
DR HSSP: P04958; IABD.
DR InterPro: IPR001064; Crystal1in.
DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 451 AA: 51823 MW: 69ABC5F030E6CD8E CRC64:

```

```

Query Match 95.3%; Score 2340; DB 2; Length 451;
Best Local Similarity 98.0%; Pred. No. 3e-131;
Matches 442; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 13 KNIDCWVNDNEDIVILKSTIILNDNDIISDISGFNSSVITTPDQOLVPGINGKAIH 72
    |||||
Db 1 KNIDCWVNDNEDIVILKSTIILNDNDIISDISGFNSSVITTPDQOLVPGINGKAIH 60
OY 73 LVNSESSEVIHVKAMDEYNMFNFVFWLVPVSAHLEOYGNNEYSIISSMKHS 132
    |||||
Db 61 LVNSESSEVIHVKAMDEYNMFNFVFWLVPVSAHLEOYGNNEYSIISSMKHS 120
OY 133 LSTSGMSVSLKGNLITWLKDSAGEVROITFRDLDPKFNAYLANKWVITITDRISSA 192
    |||||
Db 121 LSTSGMSVSLKGNLITWLKDSAGEVROITFRDLDPKFNAYLANKWVITITDRISSA 180
OY 193 NLVINGVLMGSAEITGAIREDNNITLKLDRCNNNOVYSIDKFRIFCALNKEIEKL 252
    |||||
Db 181 NLVINGVLMGSAEITGAIREDNNITLKLDRCNNNOVYSIDKFRIFCALNKEIEKL 240
OY 253 YTSIISLITFLRDFGNPLRYDTEYLLIPVASSKDVOLKNTTYMYLTNAPSYNGKLN 312

```

```

Db 241 YTSIISLITFLRDFGNPLRYDTEYLLIPVASSKDVOLKNTTYMYLTNAPSYNGKLN 300
    |||||
OY 313 YRRLVGLKFEIKRYTPNNEIDSFVKSQDEIKLYSVNNNEHIVGYPKDGNAFNNIDRI 372
    |||||
Db 301 YRRLVGLKFEIKRYTPNNEIDSFVKSQDEIKLYSVNNNEHIVGYPKDGNAFNNIDRI 360
OY 373 LRVGNAPGIPLYKKMEAVKLRLDKTYSVOLKLYDDKNASGLVGTGNGOIGNDPNRDIL 432
    |||||
Db 361 LRVGNAPGIPLYKKMEAVKLRLDKTYSVOLKLYDDKNASGLVGTGNGOIGNDPNRDIL 420
OY 433 IASNMVFNHLKDKLTLCGDWYFVPTDEGTND 463
    |||||
Db 421 IASNMVFNHLKDKLTLCGDWYFVPTDEGTND 451

```

RESULT 3

```

O08077 PRELIMINARY: PRT: 1291 AA.

```

```

AC 008077:
AC 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Botulinum neurotoxin type B (EC 3.4.24.-) (BONT/B).
DE BONT/B.
GN Botulinum botulinum.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxId=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EKLUND 17B AFCC25765;
RX MEDLINE=94122659; PubMed=7764370;
RA Hutson R.A., Collins M.D., East A.K., Thompson D.E.;
RT "Nucleotide sequence of the gene coding for non-proteolytic Clostridium botulinum type B neurotoxin: comparison with other Clostridium neurotoxins.";
RL Curr. Microbiol. 28:107-110(1994).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC ENDOPEPTIDASE THAT CLEAVES SYNAPTOSOMAL-2.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A A HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY, WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF BOTULINUM NEUROTOXIN. TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: HIGH WITH OTHER BOTULINUM NEUROTOXINS AND WITH TETANUS NEUROTOXIN.
CC -1- SIMILARITY: TO OTHER ZINC METALLOPROTEINASES IN THE ACTIVE SITE REGION.
DR EMBL: X71343; CAAS0482.1; -.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; Bontoxilysin.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27.1.
DR PRINTS: PR00760; BONTTOXILYSIN.
DR PRODOM: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR NEUROTOXIN; Transmembrane; Hydrolase; Metalloprotease; Zinc.
KW SEQUENCE 1291 AA: 150513 MW: 71BCAF23D69FAA CRC64:
SQ

```

```

Query Match 25.1%; Score 617; DB 2; Length 1291;
Best Local Similarity 31.0%; Pred. No. 1.9e-28;
Matches 149; Conservative 94; Mismatches 191; Indels 46; Gaps 14;

```

```

OY 5 TPIPSYSKNIIDCWVNDNEDIVILK-----KSTIILNDNDIISDISGFNSV 54
    |||||

```

```

Db 834 TITPDL-----TYSNIEILLIKFNKYNSEILNLIINLRYRDNLIIDLSGYGAKV 885
OY 55 TTPPAOLVPGINGK-AIHLVNNESSEVIVHKAMDIEYNMFNFTVSEFMRVPSASH 113
Db 886 EYVDGYKL-----NDKNOFKLTSSADSKIRYTOQNONTIFNSMFLDFSVSEFWIRPKYRND 941
OY 114 LQYCGNEXSIISMKKHSLSIGSGSVSLKGNLIMTLKDSAGEVQOIFR-DLPKFN 172
Db 942 IONHYNETIINCCKNN-----SGWKISIRGNMIITWTLIDINGKIKSVFEFYNIIRDIS 996
OY 173 AYLANKWVITTTINDRLSSANLYINGVLMGSAETITGALREDNNTILKDRCNNNOYV 232
Db 997 EXI-NMFEVITINN-LDNAKITINGTLESNDMDIKDIGEYIVGEGITFKLDGDVDTQFI 1054
OY 233 STDKRIFCKALNPKETIEKLYSYSLTFRLDRFGWGNPLRYDTEYLLIPVASSSKVOLCKN 292
Db 1055 WMYKYSIFMTQLOSNIKEIYKIOSYSEYKLDFWGNPLMRYNKEYYEMNAGNKNYSIKLVK 1114
OY 293 IYDYMVLTPASVTNGKLNYYRRLYNGLKFLIKRTYTPNNET-DSEYKSGDFIKLYSVN 351
Db 1115 DSSVEILLIRSKYNNONSNTYINRNLITGEKFIIRRESNOSINDDIYRKEDYTHLDLVH 1174
OY 352 NNE-HIVGPPKGNAPFNNLDRILRVGNAPGIPLYKMEAVKRLDKITYSVOLKVDCKN 410
Db 1175 HEEMRVAY-----KYFKRQEEKFLIISDSNEFYTIIEKEVDEQPTYSQOLFKKDEE 1230
OY 411 AS--LGLVGT---NGOIGNDPNRDLIASNMYFNHLKDK---ILGCDWYFVPTDEGWT 461
Db 1231 STDIGLIHMFYESGVLRRKKYKDYFCISKWLKEVKRPPKYSNIGCNWQFIPKDEGWT 1290

RESULT 4
OY 0933KO PRELIMINARY: PRT: 1291 AA.
AC 0933KO:
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxId=1491;
RN [1]
RP SEQUENCE FROM N.A.
RA Kirma N., Ferreira J.L., Baumstark B.R.;
RA STRAIN=593, AND 588;
RT "Characterization of six type A strains of Clostridium botulinum that
RT contained type B toxin gene sequences."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF300465; AAL11498.1; -.
DR EMBL: AF300465; AAL11498.1; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27; 1.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR Neurotoxin.
SQ SEQUENCE 1291 AA: 150843 MW: 7AC1737B0FA5A151 CRC64:

Query Match 24.4%; Score 600; DB 2; Length 1291;
Best Local Similarity 31.0%; Pred. No. 2e-27;
Matches 148; Conservative 96; Mismatches 193; Indels 40; Gaps 15;

OY 5 TPIPSYKNDLCWNEEDIDVILK-----KSTINDLINDNDISDISGNSVITTP 58
Db 834 TSIPF-----DLSTYNTNTLIEIFNKYNSDILNLIINLRYRDNKLIIDLSGYGAKVEYD 889
OY 59 DAOLVPGINGK-AIHLVNNESSEVIVHKAMDIEYNMFNFTVSEFWLRVPSASHLEOY 117
Db 890 GVKL-----NDKNOFKLTSSANSKIRYTOQNONTIFNSMFLDFSVSEFWIRPKYRNDGIQNT 945
OY 118 GTNEXSIISMKKHSLSIGSGSVSLKGNLIMTLKDSAGEVQOIFR-DLPKFNAYLA 176

```

```

Db 946 IHNETIINCCKNN-----SGWKISIRGNMIITWTLIDINGKIKSVFEFYSIKEDISEYI- 999
OY 177 NKWVETITINDRLSSANLYINGVLMGSAETITGALREDNNTILKDRCNNNOYSIDK 236
Db 1000 NMFEVITINN-SDNAKITINGKLESHIDIRDRIVIANDELIFKLDGNDIRQOFIMWKY 1058
OY 237 FRIECKALNPKETIEKLYSYSLTFRLDRFGWGNPLRYDTEYLLIPVASSSKVOLCKNTDY 296
Db 1059 FSIFNTELOSNIKEIYKIOSYSEYKLDFWGNPLMRYNKEYYEMNAGNKNYSIKLVK 1118
OY 297 MYLTNAPSVTNGKLNYYRRLYNGLKFLIKRTYTPNNET-DSEYKSGDFIKLYSVYNNNE 354
Db 1119 GEILRSKYNNONSNTYINRNLITGEKFIIRRSNOSINDDIYRKEDYIYLDDEFNINQEW 1178
OY 355 HIVGPPKGNAPFNNLDRILRVGNAPGIPLYKMEAVKRLDKITYSVOLKVDCKNAS-- 412
Db 1179 RVMY-----KYFKRQEEKFLIAPISDSDEFYNTIQKEVDEQPTYSQOLFKKDESTDE 1234
OY 413 LGLVGT---NGOIGNDPNRDLIASNMYFNHLKDK---ILGCDWYFVPTDEGWT 461
Db 1235 IGLIGHMFYESGVLRRKKYKDYFCISKWLKEVKRPPKYSNIGCNWQFIPKDEGWT 1290

RESULT 5
OY 092AJ8 PRELIMINARY: PRT: 1291 AA.
AC 092AJ8:
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Bont Protein.
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxId=1491;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98440323; PubMed=9767710;
RA STRAIN=CDC 3281;
RC MEDLINE=98440323; PubMed=9767710;
RA Santos-Bueiga J., Collins M.D., East A.K.;
RT "Characterization of the genes encoding the Botulinum neurotoxin
RT complex in a strain of clostridium botulinum producing type B & F
RT neurotoxins."
RL Curr. Microbiol. 37:312-318(1998).
DR EMBL: Y13630; CAA73968.1; -.
DR HSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOTOXILYSIN.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SEQUENCE 1291 AA: 150840 MW: EAD3B0E46AB2E735 CRC64:

Query Match 24.4%; Score 599; DB 2; Length 1291;
Best Local Similarity 31.0%; Pred. No. 2.3e-27;
Matches 148; Conservative 96; Mismatches 193; Indels 40; Gaps 15;

OY 5 TPIPSYKNDLCWNEEDIDVILK-----KSTINDLINDNDISDISGNSVITTP 58
Db 834 TSIPF-----DLSTYNTNTLIEIFNKYNSDILNLIINLRYRDNKLIIDLSGYGAKVEYD 889
OY 59 DAOLVPGINGK-AIHLVNNESSEVIVHKAMDIEYNMFNFTVSEFWLRVPSASHLEOY 117
Db 890 GVKL-----NDKNOFKLTSSANSKIRYTOQNONTIFNSMFLDFSVSEFWIRPKYRNDGIQNT 945
OY 118 GTNEXSIISMKKHSLSIGSGSVSLKGNLIMTLKDSAGEVQOIFR-DLPKFNAYLA 176
Db 946 IHNETIINCCKNN-----SGWKISIRGNMIITWTLIDINGKIKSVFEFYSIKEDISEYI- 999

```



```

RESULT 6
093G71
ID 093G71 PRELIMINARY; PRT: 1291 AA.
AC 093G71;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Neurotoxin type B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_Taxid=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1436;
RA Kirma N., Ferreira J.L., Baumstark B.R.;
RT "Characterization of six type A strains of Clostridium botulinum that
RT contain type B toxin gene sequences.";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
EMBL: AF295926; AA97132.1; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MPTGdase.
DR Pfam: PF01742; Peptidase_M27; 1.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 1291 AA; 150824 MW; D7CA07BAE2EBBCD2 CRC64;

```

	RESULT	
7		
045851		
ID	045851	PRELIMINARY;
AC	045851	PRT; 1268 AA.
DT	01-NOV-1996	(TREMblrel. 01, Created)
DT	01-NOV-1996	(TREMblrel. 01, Last sequence update)
DT	01-DEC-2001	(TREMblrel. 19, Last annotation update)
DE	Neurotoxin type F.	
GN	BONT / F.	
OS	Clostridium baratii.	
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;	
OC	Clostridiales; Clostridiaceae; Clostridium.	
OX	NCBI_TaxID=1561;	
OX	[1]	
RP	SEQUENCE FROM N.A.	
RP	MEDLINE=93252228; Pubmed=9486245;	
RA	Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,	
RA	Richardson P.T.;	
RT	"Nucleotide sequence of the gene coding for Clostridium baratii type F	
RT	neurotoxin: Comparison with other clostridial neurotoxins.";	
RL	FEMS Microbiol. Lett. 108:175-182(1993).	

```

QY 7 I P S Y S K N L D C W M D N E D D I D Y L K ----- K S T I L K L D I N N D I S I O F N S S V T Y P P A 60
    I I I I I : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 833 I P E L S E --- Y N D N I L H I L I R E Y K R I I D S T I L M M K E N N R F I D S S G A S I S I N G I 888
    I I I I I : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 Q L P G I N G K A I H L Y N N S S E V I V H K A M D I E Y N D M F N F T V S F W L R Y P K V S A H L E O Y G T N 120
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 889 Y I I - S T I N R O F O G I Y S S R L S E V N I T Q N T I I T Y N S R Y Q N F S V F W R I P K - - N I L K M - N N 944
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 E Y S I I S M K K H S I S G S V S I L K G N N L W T L K D S G E V R Q I T E R ----- D L P D K R N A V L 175
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 945 E Y T I I N C M R N N N --- S G W K I S L N Y N N I I W T L O D T T G N N O K L E V N Y T Q M I D I S D --- Y I 996
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 176 A N K V E T I I T I N D R L S A N L Y I N G V L W G S A I T I G A I R E D N I N T L K L D R C N N N O V S I D 235
    I I I : : : : : I : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 997 - N K W T E V T I I N N R I G H S K I I N G L U D O K S I L M G I N H V D N I L F R K I Y V C - N O T R Y V G I R 1054
    I I I : : : : : I : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 236 K E R F E C K A I M P K E I E K L Y T S I T T E R D W G N P L R Y D E Y Y L I --- P V A S S K D V O L K 291
    I : I I : : : : : I : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1055 Y E F L E F N E L D K I E T I E T I L Y H S E P D S T I L K D W G N V L Y K N K Y Y L L N I L L K P M M S Y T K S I D 1114
    I : I I : : : : : I : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 292 N I T D Y M V L T N A P S Y T N G K L N I Y Y R - R L Y N G L K F I K A Y --- T P N N E I D S F P K S O D F I K L Y 348
    I I : : : : : I : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1115 N I --- N R Q R I G S K T I F S N A R L Y T G E V I L I R K V S T D T S I N D N E F V R K M D Y T I Y N 1167
    I : : : : : I : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 349 S Y N N N E H I V G I P K C N A F N --- N I D R L R Y G V A N P G I P L Y K K M E A V L R L K I T Y S V Q L T 405

```


QY	442	LKDKI--LGCWYFYPTEG	460
Db	1232	MRDNTNSNGFWEFTSEHG	1252
RESULT	11		
ID	057236	PRELIMINARY;	PRT: 1278 AA.
AC	057236; Q45863;		
DT	01-NOV-1996 (TREMblrel. 01, Created)		
DT	01-NOV-1996 (TREMblrel. 01, Last sequence update)		
DT	01-JUN-2001 (TREMblrel. 17, Last annotation update)		
DE	Bacillus neurotoxin type F (BONT/F protein).		
OS	Clostridium botulinum.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;		
OC	Clostridiales; Clostridiaceae; Clostridium.		
ON	NCBI_taxid=1491;		
RX	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NCCTC 10281;		
RL	Hutson R.A., Collins M.D.;		
RL	Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Elmore M.J., Bodsworth N.J., Whelan S.M., Minton N.P.;		
RL	Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE OF 635-1000 FROM N.A.		
RC	STRAIN=NCCTC 1028;		
RC	MEDLINE=94013372; PubMed=8408542;		
RA	Campbell K., East A.K., Collins M.D.;		
RT	"Gene probes for identification of the botulin neurotoxin gene and		
RL	specific identification of neurotoxin types B, E, and F.;"		
RN	J. Clin. Microbiol. 31:2255-2262(1993).		
RN	[4]		
RP	SEQUENCE OF 1-27 FROM N.A.		
RC	STRAIN=LANGE LAND;		
RC	MEDLINE=98404102; PubMed=9732534;		
RA	East A.K., Bhandari M., Helm S., Collins M.D.;		
RT	"Analysis of the botulin neurotoxin type F gene clusters in		
RT	proteolytic and nonproteolytic Clostridium botulinum and Clostridium		
RT	baratii.;"		
RL	Curr. Microbiol. 37:262-268(1998).		
DR	EMBL; X81714; CAA57358.1; -;		
DR	EMBL; L35496; AAA23210.1; -;		
DR	EMBL; X70821; CAA50152.1; -;		
DR	EMBL; X99064; CAA67512.1; -;		
DR	HSSP; P10845; 3B7A.		
DR	MEROPS; M27.002; -;		
DR	InterPro; IPR000395; Bontoxilysin.		
DR	InterPro; IPR000130; Zn_MTPeptide.		
DR	pfam; PF01742; Peptidase_M27; 1.		
DR	PRINTS; PR00760; BONTTOXILYSIN.		
DR	ProDom; PD001963; Bontoxilysin; 1.		
DR	PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.		
DR	Neurotoxin.		
SQ	SEQUENCE 1278 AA; 147073 MW; A1BFL318431D6918 CRC64;		
Query Match	22.5%;	Score 553;	DB 2; Length 1278;
Best Local Similarity	30.0%;	Pred. No. 1.2e-24;	
Matches 138; Conservative	92;	Mismatches 152;	Indels 78; Gaps 18;
QY	29	LKSTLLNDINNDIISDTSFGNSSVITTPDAQVLPGINGKAIHLVNNSSSEVYHKAMD	88
Db	866	IKDNDILDMRYENMKNFIDISGYSINSGVDYIT-STRNQGFGYSSKPSKPSVNTAONND	924
QY	89	IEYNDMEFTVSEMLRYPKYSASHLEQYG-TNEYSIISMKKHSLSIGSGVSATKCN	147
Db	925	IYNGRNYONFSISFWRIK---YFNKYNLNNETITIDICIRNN-----SGKISILNTNK	976
QY	148	LWTLKDSAGEVROITFR-----DLPDKFNAYLANKWVFTITNDRLSSANLYINGVLNG	202

[illegible]


```

Db 1059 LTRDNPITFSKELSNEDINIVEGQILRNVIKDYWGPNPLKFDTEYYII--NDNYIDRYIA 1116
QY 292 NITDWMYLTNAPSYTNGKINITYRRLYNGLKFLIKRTPNNEIDSEFKSGDFIKLYSYN 351
Db 1117 PKSNILVLVOYFDRS-----KLYTGNPTITIKSVSDKNPY--SRILNDNIMFHLN 1166
QY 352 NNEHIVGYPRKDGNAFNNLDRIIRVGNAPGIPLYKK-MEAVKLR-DLKTYSVOLKLYDDK 409
Db 1167 SKRYMI-----IRDTDTI-----YAIEGRECSKNCVYALKQSNIGNYGI--GIPIK 1212
QY 410 NASLGLVGTN--GOIGNDPNRDIILASN--WYFN-----HLKDKILGCD-- 450
Db 1213 N----IVSQNKYCSQIFSSFMKNTMLLADIKPWRFSFENATYPVAVTNYETKLLSTSSF 1268
QY 451 WYFVPTDEGW 460
Db 1269 WKFIIRDPGW 1278

```

RESULT 15

```

Q45849 ID 045849 PRELIMINARY: PRT: 1280 AA.
AC 045849:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Neurotoxin.
OS Clostridium botulinum C.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=36628;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6813;
RX MEDLINE=96156810; PubMed=8593068;
RA Moritski K., Koura M., Fujii N., Fujinaga Y., Inoue K., Syuto B.,
RA Oguma K.;
RT "Molecular cloning of the gene encoding the mosaic neurotoxin,
RT composed of parts of botulinum neurotoxin types C1 and D, and PCR
RT detection of this gene from Clostridium botulinum type C organisms.";
RL Appl. Environ. Microbiol. 62:662-667(1996).
DR EMBL: D49440; BAC08418.1; -.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_Metpeptidase.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOXILYSIN.
DR PRODOM: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SO SEQUENCE 1280 AA; 147815 MW; 74F813B228B8C989 CRC64;

```

Query Match 17.0%; Score 417.5; DB 2; Length 1280;
 Best Local Similarity 26.1%; Pred. No. 1.3e-16;
 Matches 128; Conservative 100; Mismatches 179; Indels 83; Gaps 20;

```

QY 3 ESTPIFSSYKMLDCWVNEEDIVY-----LKRSTIINLDINNDIISDGFNSSVIT 56
Db 840 FENLIPF-----NIFSYTNLSLKDIIINEFNSINDSKILSLQKKKNAALVDTSGYNAEVL 895
Db 896 EDDVGV-----NTIYTNDFKLSSSGDKTIIVNNIILYSALYENSVSFWIKISKDLTN 949
QY 57 YPDADLVGICNGKAIHL-----VNESSEVIYVHRAMDIENDMFNFTVSFWLVRPK-VSA 111
Db 896 EDDVGV-----NTIYTNDFKLSSSGDKTIIVNNIILYSALYENSVSFWIKISKDLTN 949
QY 112 SHLEQYGTNEYYSIISMKHSLSTSGWVSLSKGNLITLTKDSAGEVROITFRDLDPKR 171
Db 950 SH-----NETIISINQON-----SGWKLCLIRNGNIEIWIQDINRKYSLLIFDYSESLS 998
QY 172 NAYLANKWVFTITNDRLSALYINGVLMGSAEITGLAIREDNNTITLDRCNNNOY 231
Db 999 HTGYTNKKWFYITITNINMGYMKLYINGELKQSERIEDLNEVAKLDKTIIVGIDENIDENOM 1058

```

```

QY 232 VSIDKRIEFCALNPKETIEKLYTSTLITFLRDFWGNPLRYDTEYYLIPASSKDVOLK 291
Db 1059 LTRDNPITFSKELSNEDINIVEGQILRNVIKDYWGPNPLKFDTEYYII--NDNYIDRYIA 1116
QY 292 NITDWMYLTNAPSYTNGKINITYRRLYNGLKFLIKRTPNNEIDSEFKSGDFIKLYSYN 351
Db 1117 PKSNILVLVOYFDRS-----KLYTGNPTITIKSVSDKNPY--SRILNDNIMFHLN 1166
QY 352 NNEHIVGYPRKDGNAFNNLDRIIRVGNAPGIPLYKK-MEAVKLR-DLKTYSVOLKLYDDK 409
Db 1167 SKRYMI-----IRDTDTI-----YAIEGRECSKNCVYALKQSNIGNYGI--GIPIK 1212
QY 410 NASLGLVGTN--GOIGNDPNRDIILASN--WYFN-----HLKDKILGCD-- 450
Db 1213 N----IVSQNKYCSQIFSSFMKNTMLLADIKPWRFSFENATYPVAVTNYETKLLSTSSF 1268
QY 451 WYFVPTDEGW 460
Db 1269 WKFIIRDPGW 1278

```

Search completed: December 1, 2002, 11:23:47
 Job time : 90 secs

